

Query Match	100.0%;	Score 1527;	DB 6;	Length 1527;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1527;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;



Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CTCATGCACTCGCGCTTCGGCTTCCTCCAGTGTGGTTCGTTTCAACGCAACCAAG	60	
Db	1	CTCATGCACTCGCGCTTCGGCTTCCTCCAGTGTGGTTCGTTTCAACGCAACCAAG	60	
QY	61	CGTTTCGGGGCAATGAGTTCTTGGGCGCGCGTATTTGGTATGCGGATGTTCCCGAG	120	
Db	61	CGTTTCGGGGCAATGAGTTCTTGGGCGCGCGTATTTGGTATGCGGATGTTCCCGAG	120	
QY	121	CTTGGTGAACGGCTACGAGTGGCGCGCAACATGCTCGGGGCAATGCAATGCTGTC	180	
Db	121	CTTGGTGAACGGCTACGAGTGGCGCGCAACATGCTCGGGGCAATGCAATGCTGTC	180	
QY	181	CCTGTTTGGTTTATAGTTTGGCCAAAGCCGTTTACCAAGGCAACGCTTCTCTGCTGGT	240	
Db	181	CCTGTTTGGTTTATAGTTTGGCCAAAGCCGTTTACCAAGGCAACGCTTCTCTGCTGGT	240	
QY	241	GTTTCTTGGATTCCTGGCAACGATCGAGAGTTCTTGCACAAGCGACTCAAGGGCACTGC	300	
Db	241	GTTTCTTGGATTCCTGGCAACGATCGAGAGTTCTTGCACAAGCGACTCAAGGGCACTGC	300	
QY	301	AGACTTCTGATCAGTCCAGTGTGAGCTGCTGCTCAGCGGATTCCTTACATTCATCGC	360	
Db	301	AGACTTCTGATCAGTCCAGTGTGAGCTGCTGCTCAGCGGATTCCTTACATTCATCGC	360	
QY	361	CATTGGCCCAAGCAATCGCTGGGTGGCGGATGCTGGCAACGCTTCAAGGCACTTTA	420	
Db	361	CATTGGCCCAAGCAATCGCTGGGTGGCGGATGCTGGCAACGCTTCAAGGCACTTTA	420	
QY	421	TGATTTGGTGTGCTAGTGGCGGTTGCTCTTGGTCTGCTTACTCACAATCGTCAT	480	
Db	421	TGATTTGGTGTGCTAGTGGCGGTTGCTCTTGGTCTGCTTACTCACAATCGTCAT	480	
QY	481	CATGCTGTGCAACCACTCTTCCGCAATGAGCTGGAGCTGTTTAAACAGGCTGATC	540	
Db	481	CATGCTGTGCAACCACTCTTCCGCAATGAGCTGGAGCTGTTTAAACAGGCTGATC	540	
QY	541	CTTCATCTTCGCAACCGCACTATGCTAATATCGCCAGGGTGGCGATGTTTGGCAGT	600	
Db	541	CTTCATCTTCGCAACCGCACTATGCTAATATCGCCAGGGTGGCGATGTTTGGCAGT	600	
QY	601	GTTTCTTCTGGCAAGAGTGAAGCTCAAGGCTTTCGAGGTGCTTCAAGTGTCTCCGC	660	
Db	601	GTTTCTTCTGGCAAGAGTGAAGCTCAAGGCTTTCGAGGTGCTTCAAGTGTCTCCGC	660	
QY	661	TGTTCTTGTGATTAACGAGCTCGATCTTCGCTGTGGAACCTTCGCTGCGCTGGCGGTT	720	
Db	661	TGTTCTTGTGATTAACGAGCTCGATCTTCGCTGTGGAACCTTCGCTGCGCTGGCGGTT	720	
QY	721	CTTCATCGGTATCGGTACCGCACTATCGGTGGCGCTTTGATTGCACTCTTTAATATCAA	780	
Db	721	CTTCATCGGTATCGGTACCGCACTATCGGTGGCGCTTTGATTGCACTCTTTAATATCAA	780	
QY	781	GGCAGTTGGTGGCGCTGCAAGTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840	
Db	781	GGCAGTTGGTGGCGCTGCAAGTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840	
QY	841	GGTCATGTTCTGTTGTCAGTGTGATCTTCTCATCGCATTCGGCGCAGCGATTGC	900	
Db	841	GGTCATGTTCTGTTGTCAGTGTGATCTTCTCATCGCATTCGGCGCAGCGATTGC	900	
QY	901	TTATGGCTTTTACTTGGTTCGCGCAACCGCAGCATTTGATCCAGATGCAACCGCTCTCC	960	
Db	901	TTATGGCTTTTACTTGGTTCGCGCAACCGCAGCATTTGATCCAGATGCAACCGCTCTCC	960	
QY	961	AGTGCCTGCAAGGAACCAACCAAGCAGAGCAACCGCAGAGATTTTCAAACGATTTC	1020	
Db	961	AGTGCCTGCAAGGAACCAACCAAGCAGAGCAACCGCAGAGATTTTCAAACGATTTC	1020	
QY	1021	CACCATCATCCAGGCACTTTTGAACGGTGAAGCTATTGCACTGAGCAGCTGACGATGC	1080	
Db	1021	CACCATCATCCAGGCACTTTTGAACGGTGAAGCTATTGCACTGAGCAGCTGACGATGC	1080	

QY	1081	CATGTTTGCAGCGGAAAGCTTGGCTCGGCGCTTGCCATCGTCCCAACCAAGGGCAGTT	1140	
Db	1081	CATGTTTGCAGCGGAAAGCTTGGCTCGGCGCTTGCCATCGTCCCAACCAAGGGCAGTT	1140	
QY	1141	AGTTTCTCCGGTGAAGTGGTGGGCAATTCCTCCATTCCTCCATCTGGCCATGCTTTCGAGT	1200	
Db	1141	AGTTTCTCCGGTGAAGTGGTGGGCAATTCCTCCATTCCTCCATCTGGCCATGCTTTCGAGT	1200	
QY	1201	TGCGACCAAGGCTGAGATGGTTCGAATGTGGATATCTTGATGACATTCGTTTCGACAC	1260	
Db	1201	TGCGACCAAGGCTGAGATGGTTCGAATGTGGATATCTTGATGACATTCGTTTCGACAC	1260	
QY	1261	AGTAAACCTCAACCGGCAACGCTTTAAACCGCTGGAAGAGCAGGCGATGAAAGTCAAAGC	1320	
Db	1261	AGTAAACCTCAACCGGCAACGCTTTAAACCGCTGGAAGAGCAGGCGATGAAAGTCAAAGC	1320	
QY	1321	AGGGAGCTGCTGTGTAATTCGATATTCGATTCATTAAGGCTGCAGGTTATGAGGTAAC	1380	
Db	1321	AGGGAGCTGCTGTGTAATTCGATATTCGATTCATTAAGGCTGCAGGTTATGAGGTAAC	1380	
QY	1381	CAGCCGATGTGTTTTCGAATTTACAGAAACCGGACCTGTAACACTTACGTTTGGG	1440	
Db	1381	CAGCCGATGTGTTTTCGAATTTACAGAAACCGGACCTGTAACACTTACGTTTGGG	1440	
QY	1441	CGAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAGCGGTGCAGCAAC	1500	
Db	1441	CGAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAGCGGTGCAGCAAC	1500	
QY	1501	ACCATTAAGTTGAAACCTTGAGTGTTCG	1527	
Db	1501	ACCATTAAGTTGAAACCTTGAGTGTTCG	1527	

RESULT 3  
AP005282/c

LOCUS  
DEFINITION  
AP005282 320550 bp DNA linear BCT 08-AUG-2002  
Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 9/10.

ACCESSION  
VERSION  
AP005282 BA000036  
AP005282.1 GI:21325287

KEYWORDS  
SOURCE  
ORGANISM

Corynebacterium glutamicum ATCC 13032  
Corynebacterium glutamicum ATCC 13032  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

## REFERENCE

1  
Nakagawa, S.  
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032  
Unpublished  
2 (bases 1 to 320550)  
Nakagawa, S.  
Direct Submission  
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hako Kogyo Co.  
Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,  
Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com,  
Tel: 81-44-829-3031, Fax: 81-44-813-1651)  
This sequence is conducted by collaboration of Kyowa Hako Kogyo  
Co. Ltd. And Kitasato University.

## COMMENT

FEATURES  
source

1. 320550  
/organism="Corynebacterium glutamicum ATCC 13032"  
/mol\_type="genomic DNA"  
/strain="ATCC 13032"  
/db\_xref="taxon:196627"  
/note="ATCC 13032"  
188. .814  
/gene="Cgl2516"  
188. .814  
/gene="Cgl2516"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical protein"

```
/protein id="BAB99909.1"
/db_xref="GI:21325289"
/translation="MYRAPDFVPLSMIVLVYTRKDRMETLAQAQTILEKWCVAP
THASFVSIKAIPILSLITLIVTVNGISSGNVPQPPALBQVTRDVVNKNIYERLKL
GLVISPHELTAAQTIAQRNADSEKRPDPENLVLIQONLPYANADTIIVDR
FLNSPDHVKLLANDYEAIGVAVYKGDHAWIVVEFTVAPADSVESBTE"
811..2151
/gene="Cg12517"
CDS
811..2151
/gene="Cg12517"
/codon_start=1
/transl_table=11
/product="Nicotinic acid phosphoribosyltransferase"
/protein id="BAB99910.1"
/db_xref="GI:21325289"
/translation="NNTMPSEFSSNRSTALLTDKVELTWLQALADGSAERPSPEVF
SRRLPNERGYAGTAVTAKAIDRFVTEBQADLDFDORTLEYLRNRYPTQGVGD
YREGIYFQPSLLVTRGTFAVCVILETVIISINADAVASAAAMVTAADGREIIE
MGSRTHEYSAVTAAGYVLAGFTTSLEAARYGIPASGTSAAHATLLHNDGTP
NEAAFKQAVESLGVDTLLVDYDI TOGVATAIRVAGPDIGVRI DSGDGLVARKV
RKQLDLNHNKTIWSSDLDEFALAGRGEPDVFGVTSVTGSGAPTAGLVYKIV
EVAGHPVAKRSKSYGGKKAIVHRKSGTAIEI VIPFNAEPDTFGKDLISLTI
PLMRDGETVPEGLTLEDGRAYLAKQLVSLPWEGLASRDEPVIHTRFVGFPPAA"
2438..2746
/gene="Cg12518"
CDS
2438..2746
/gene="Cg12518"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein id="BAB99911.1"
/db_xref="GI:21325290"
/translation="MGLRLNLPGRPPQKTOITWKLPAQNRPOAPLSKOSNOTVNRCEFA
WLSSTPWLDAQCLRLNHPQIPQIRSMRPKSCLLILPKRWAVPVVLDKRWLRL"
2866..4608
/gene="Cg12519"
CDS
2866..4608
/gene="Cg12519"
/codon_start=1
/transl_table=11
/product="Rad3-related DNA helicases"
/protein id="BAB99912.1"
/db_xref="GI:21325291"
/translation="MSTATIALQROLVNERDLRLVDALEPLMERPTFAIMKGRSNYL
CMKVVARQELNQBALIEQDI SWLKGHIVELNWEANETETGDRDLDPGYPDLANK
QVSVTARECIGASCPHGEDCFATARGAKKADVVVTHALLAIDALSDSVLPEHD
VVVIDEAHLDGRIITAVASAEITVNSLNLAARRAKLSDKREERVQETAGDLETLLQ
TWQGRWMDSGSGTILVALDALWALRAQIAGAPEGAANDPERFAERQNSHLIM
ETHDANVRILEVFABEDPSKQDVVWVHNHDDRGSINLVAPLSVAGLHKEKLFANV
VLASATLTIGGNFNAMASWGLPKGSWDSMDAGTPDPAKSGILYATRHLPDGRDGL
PRETLEIYELLITAGGRTLGLFSSKRAAEQATKAMRLLEPFDVLCQDDNTAALVKK
FSDSENTCLFGLTLWQGVDPGRSLSLVLDRIIFPPRPDDELQARKRAADAEGRNG
FMEVAATHAALMAQAGELLRHVGDRGVAVLDHRLSTKRYGGFLRFSNRPFWETTN
PPTVRAALKRLVTK"
complement(4618..5511)
/gene="Cg12520"
CDS
complement(4618..5511)
/gene="Cg12520"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein id="BAB99913.1"
/db_xref="GI:21325292"
/translation="MISPYGSGVQGDNNNNNSQGNPSPGQNGYQGYQAPYQGYQYQ
PFQGFNAYSSPIPEVPQSQMQEAWRSFDLGTVPQAKGFTATWQAWLSALIYF
AVLLVMEFLMPLVSVLAATSSGSDSAAIATGTSFFGFMIMVIAFISFVSLNC
YRNAARVGEQITIQSPFKMGLKALGIYILINIVIFIGMILIPGIIAAVLIF
AVPVAFLRDAGISGAFSAWKAWSKNVQGVILLELAIFALSFLGSAVILGMVTTPL
TLLVAYAFQTAGGPIMQORQ"
complement(5587..6312)
/gene="Cg12521"
CDS
complement(5587..6312)

/gene="Cg12521"
/translation="Domain of unknown function UPF0099"
/codon_start=1
/transl_table=11
/product="Uncharacterized ACR"
/protein id="BAB99914.1"
/db_xref="GI:21325293"
/translation="MPKSLDLAHLKARASDGDSPSDPSTVQAMQIALHLPKQ
NPPRTDYLEAARSVKLCIDERVSTDPDFAALRWYGHILRKVSRRNAAWDRV
QDLPGVTVEDDAAKVRAFLPSAVDVDPADIKKLQISCTELPDEPNINDEVPVLIYD
ESLKWTLGAAQVGHASMLLAHQPFWEVQWEAADFALHREIPSEFLALIESPG
AVPRDGGFTVAPNTVTVAIP"
complement(6305..7324)
/gene="Cg12522"
CDS
complement(6305..7324)
/gene="Cg12522"
/EC_number="3.1.3.3"
/translation="haloacid dehalogenase-like hydrolase"
/translation="TIGR00099: conserved hypothetical protein"
/translation="TIGR00338: serB: phosphoserine phosphatase SerB"
/codon_start=1
/transl_table=11
/product="Phosphoserine phosphatase"
/protein id="BAB99915.1"
/db_xref="GI:21325294"
/translation="MHGQSVVVELQETVQSSRPSRVVVLGDPVDALDISRIGTTL
ADYDANIDTIRGISDPYVTVLELVKVTVDVSPGGEAMKALAAATSELNVDAIERS
GILRSKRLVCFDCDSTILTGETVIEMLAHAGKEAAVATERAMRGELDBESLRER
VKALAGLSDAVIDEVAALLETPTGARTTIRINRNGYQTAVVSGGFIQVLEGLAELE
LDYVRANTLEIVDGKLTGNVTGKIVDRAAKAEFLREFAADSLGKMWQTVAVGDGANDI
DMLSAAGLVAFNAKPAKLEADTSVNHPLFDELVHIMGISRDEIDLDAQEDGTTTHRV
PLTNA"
complement(7811..9565)
/gene="Cg12523"
CDS
complement(7811..9565)
/gene="Cg12523"
/EC_number="1.9.3.1"
/translation="Cytochrome C and Quinol oxidase polypeptide"
/codon_start=1
/transl_table=11
/product="Heme/copper-type cytochrome/quinol oxidases, subunit 1"
/protein id="BAB99916.1"
/db_xref="GI:21325295"
/translation="MTAVAPRVGDHVAQRPBPTGHARGSKAWLMMTTTDHKLQIGM
YIIMSPSFYFGLGIMALLIRALFTPGIQLFSLNEQNFQI.FTHGFMVLLLYXTPIVWG
FANYVLPQIGADPVAFPRLNAGFWITTVGVAMLTGFTPGGAADFQWMTSPUSD
AIHSPGLGSDMWIVGVGATIGSVASAINMLTTLILRAPGTMFRMP.LFTWNIFVVS
VIALILFLLLAALGVLYDRKLGHLDPANGSLLOHLEFWFQGHPPVYVLAIPFF
GIVSIIIPVSRKPMFVYGLI.FATLISGALSMVAHMFVTVGAVLLPFPFMTFLI
SVPTGVKPFNVGTWKGHITWETPMKSVGMATFLGGLTGIMLASPLDPLHLADS
VYLLAHFHTYLTGTVFASCAGYVFPWMTGEMMDERLGT.FHWT.FYVGHGTFLIQ
HWVGNMPPRRYADYLDSDGFTLYNQISTVTFSLGLSLVIFINWVFKSWRGTBELTV
DDPWYGNLSLEWATSPPPPHNFASLPIRISRPAPFELHYPHMIERMAEAHTGHDD
INAPBLGTAPALADSSR"
9701..10009
/gene="Cg12524"
CDS
9701..10009
/gene="Cg12524"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein id="BAB99917.1"
/db_xref="GI:21325296"
/translation="MPKKNVFTIHEVAARFSLHVAFTNQRCEGVHNLTDISPGSPT
PTVNECHRFHNSPRGSIEWLRPFQGRFLLRDVPFPPRGANRSPTPICASEKPAF"
complement(10086..11090)
/gene="Cg12525"
CDS
complement(10086..11090)
/gene="Cg12525"
/translation="Ribonucleotide reductase, small chain"
/translation="Ribonucleotide reductase, small chain"
```

```
/codon_start=1
/transl_table=11
/product="Ribonucleotide reductase beta subunit"
/protein_id="BAB99918.1"
/db_xref="GI:21325297"
/translation="MAADSDLSVHDHAYLKEHVAPVKAINWNSIPDSKDLVMDRLGN
FWLPEKVPVNSDIKSRITNEVEKAAATMRVFTGLTLLDTIQGTGVAISLLPADSLHE
EALVNTIAFMESVHAKSYNIFMTLASTAEIINDAFRWSBENENLQKAKIILSYBGD
DPLRKIASVILLESFLYSGFVLPMTWSHSLKNTADVRLIIRDEAVHGYIVKY
QKAVAKETPEROBELEKEYTFDLYLDYDNETOYSEDLVLDLGTWEDVKEFLRYNANKA
LNNLVEGELFPADETKVSNILSALSFNADENHDFPSGSSSYVIGKAENTEDDDWDF
"
gene 11555..12043
/genes="Cgl12526"
CDS
11555..12043
/genes="Cgl12526"
/note="PF00210: Ferritin

Query Match          99.2%; Score 1515; DB 1; Length 320550;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1526; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CTTATGCGATCTGCGCGCTTCGGTCTTCGCCAGTGTGGTTCAGCGCAACCAAG 60
DB 148077 CTTATGCGATCTGCGCGCTTCGGTCTTCGCCAGTGTGGTTCAGCGCAACCAAG 148018

QY 61 CGTTTCGGCGGCAATGAGTTCTTGGCGCGCGCGTATTGGTATGGCGATGGTGTCCCGAG 120
DB 148017 CGTTTCGGCGGCAATGAGTTCTTGGCGCGCG -GTATTGGTATGGCGATGGTGTCCCGAG 147959

QY 121 CTTTGTGAACGGCTACGACGTGGCGCGCACCATGCTGGTGGCGGCGAAATGCCAATGTGGTC 180
DB 147958 CTTTGTGAACGGCTACGACGTGGCGCGCACCATGCTGGTGGCGGCGAAATGCCAATGTGGTC 147899

QY 181 CCGTTTCTGTTTAGATGTTGCCAACCGGTTTACGAGGCGACCGTGTCTCTGTCTGTT 240
DB 147898 CCGTTTCTGTTTAGATGTTGCCAACCGGTTTACGAGGCGACCGTGTCTCTGTCTGTT 147839

QY 241 GGTTCCTTGATCTTCGGCAACGATCGAGAAGTTCTCTGCAACGCACTCAAGGGCACTGC 300
DB 147838 GGTTCCTTGATCTTCGGCAACGATCGAGAAGTTCTCTGCAACGCACTCAAGGGCACTGC 147779

QY 301 AGACTTCCTGATCACTCCAGTGTGACGTTGTGCTCACCGGATTCCTTACATTCATCGC 360
DB 147778 AGACTTCCTGATCACTCCAGTGTGACGTTGTGCTCACCGGATTCCTTACATTCATCGC 147719

QY 361 CATTCGCCACGACATCGTGGTGGCGGATGTGCTGGCACAGGCTACAGGGACTTTA 420
DB 147718 CATTCGCCACGACATCGTGGTGGCGGATGTGCTGGCACAGGCTACAGGGACTTTA 147659

QY 421 TGATTTCGGTGTGCCAGTCCGGCGGTCTGCTCTTCGGTCTGCTTACTCAACCAATCGTCAT 480
DB 147658 TGATTTCGGTGTGCCAGTCCGGCGGTCTGCTCTTCGGTCTGCTTACTCAACCAATCGTCAT 147599

QY 481 CACTGTGCTGCACAGTCCTTCCCGCAATGAGCTGTGAGCTGTGAGCTTTAACAGGGTGGATC 540
DB 147598 CACTGTGCTGCACAGTCCTTCCCGCAATGAGCTGTGAGCTGTGAGCTTTAACAGGGTGGATC 147539

QY 541 CTTTCATCTTCGCAACGGGATCTATGCTATATATCGCCACAGGCTGCGGATGTTTGGCAGT 600
DB 147538 CTTTCATCTTCGCAACGGGATCTATGCTATATATCGCCACAGGCTGCGGATGTTTGGCAGT 147479

QY 601 GTTCTTCTCGCGCAAGAGTGAAGGCTCAAGGCGCTTGCAGGCTTCAGGCTGTCCCG 660
DB 147478 GTTCTTCTCGCGCAAGAGTGAAGGCTCAAGGCGCTTGCAGGCTTCAGGCTGTCCCG 147419

QY 661 TGTCTTGTGATTAACGGAGCTCCGATCTTTCGGTGTGAACCTTCGCGTCCGCTGGCGGTT 720
DB 147418 TGTCTTGTGATTAACGGAGCTCCGATCTTTCGGTGTGAACCTTCGCGTCCGCTGGCGGTT 147359

QY 721 CTTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATATCAA 780
DB 147358 CTTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATATCAA 147299
```

```
QY 781 GGCAGTTGCGTGGCGCTGCAGGTTTCTTGGGTGTGTGTTCTATTGATGCTCCAGATAT 840
DB 147298 GGCAGTTGCGTGGCGCTGCAGGTTTCTTGGGTGTGTGTTCTATTGATGCTCCAGATAT 147239

QY 841 GGTTCATGTTCTTGGTGTGTCAGTGTGTTACCTTCTTTCATGCAATTCGGCGCAGGATTGC 900
DB 147238 GGTTCATGTTCTTGGTGTGTCAGTGTGTTACCTTCTTTCATGCAATTCGGCGCAGGATTGC 147179

QY 901 TTATGGCCTTTACTTGGTTCGCGGCAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCC 960
DB 147178 TTATGGCCTTTACTTGGTTCGCGGCAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCC 147119

QY 961 AGTGCCCTGCAGGAACACCAAGCCGAGCAGAAAGCACCGCAGAAATTTTCAACCAATTC 1020
DB 147118 AGTGCCCTGCAGGAACACCAAGCCGAGCAGAAAGCACCGCAGAAATTTTCAACCAATTC 147059

QY 1021 CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080
DB 147058 CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 146999

QY 1081 CATGTTTGCAGCGGAAAGCTTGGCTCGGGCGGTGGCCATCGTCCCAACCAAGGGGCAATT 1140
DB 146998 CATGTTTGCAGCGGAAAGCTTGGCTCGGGCGGTGGCCATCGTCCCAACCAAGGGGCAATT 146939

QY 1141 AGTTTCTCCGGTGAGTGGAAAGATTGCTGTGGCATTCCTCATCTGGCCATGCTTTCCGAGT 1200
DB 146938 AGTTTCTCCGGTGAGTGGAAAGATTGCTGTGGCATTCCTCATCTGGCCATGCTTTCCGAGT 146879

QY 1201 TGCACCAAGGCTGAGGATGGTTTCCAATGTGATATCTTGAATGACATTCATTGTTTCCACAC 1260
DB 146878 TGCACCAAGGCTGAGGATGGTTTCCAATGTGATATCTTGAATGACATTCATTGTTTCCACAC 146819

QY 1261 AGTAACTCTCAACGGCACGCACTTTAAACCGCTGGAAGACGAGGCGATGAAGTCAAAGC 1320
DB 146818 AGTAACTCTCAACGGCACGCACTTTAAACCGCTGGAAGACGAGGCGATGAAGTCAAAGC 146759

QY 1321 AGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAGGCTGAGGTTATCAGGTTAAC 1380
DB 146758 AGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAGGCTGAGGTTATCAGGTTAAC 146699

QY 1381 CAGCGCGATTTGTTTTCGAATTAAGAAACCGGACCTGTAAACACTTACCGTTTGGG 1440
DB 146698 CAGCGCGATTTGTTTTCGAATTAAGAAACCGGACCTGTAAACACTTACCGTTTGGG 146639

QY 1441 CGAAATTTGAAGCGGAGCGCAACCTGCTCAAAGTGAAGAAAGAAAGCGGTGCCAGCAAC 1500
DB 146638 CGAAATTTGAAGCGGAGCGCAACCTGCTCAAAGTGAAGAAAGAAAGCGGTGCCAGCAAC 146579

QY 1501 ACCATAAGTTGAACCTTCAGTGTTCG 1527
DB 146578 ACCATAAGTTGAACCTTCAGTGTTCG 146552
```

```
RESULT 4
BX927155/c
LOCUS BX927155.1
DEFINITION Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5, complete genome; segment 8/10.
ACCESSION BX927155
VERSION BX927155.1
KEYWORDS complete genome.
SOURCE Corynebacterium glutamicum ATCC 13032
ORGANISM Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 349136)
AUTHORS Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkowski, A., Dusch, N., Eggeling, L., Eikmanns, B. J., Gaigalat, L., Goessmann, A., Hartmann, M., Hutmacher, K., Kramer, R., Linke, B., McHardy, A. C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A., Rey, D. A., Ruckert, C., Rupp, O., Sahn, H., Wendisch, V. F., Wiegand, I. and Tauch, A.
```

TITLE The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids

JOURNAL J. Biotechnol. 104 (1-3), 5-25 (2003)

MEDLINE 22830012

PUBMED 12948626

REFERENCE 2 (bases 1 to 349136)

AUTHORS Kalinowski, J

TITLE Direct Submission

JOURNAL Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer Genomforschung, Universitaet Bielefeld; Universitaetsstrasse 25, 33615 Bielefeld, Germany

COMMENT E-mail:Joern.Kalinowski@cebiotec.uni-bielefeld.de

This sequence was accomplished by collaboration between Degussa AG and Bielefeld University.

Join(BX927148.1:1..348071,BX927149.1:51..349887, BX927150.1:51..348475, BX927151.1:51..349459, BX927152.1:51..349799, BX927153.1:51..349584, BX927154.1:51..349575, BX927155.1:51..349136, BX927156.1:51..349115, BX927157.1:51..140057)

FEATURES

Location/Qualifiers

1..349136

/organism="Corynebacterium glutamicum ATCC 13032"

/mol\_type="genomic DNA"

/strain="DSM 20300 = ATCC 13032"

/db\_xref="taxon:196627"

/note="IS fingerprint type: 4-5"

complement(67..729)

/locus\_tag="cg2562"

complement(67..729)

/locus\_tag="cg2562"

/codon\_start=1

/transl\_table=11

/product="Serine/threonine specific protein phosphatase"

/protein\_id="CAF20997.1"

/db\_xref="GI:41326515"

/translation="WYKQLDSLVFTDGESIAKARLASNTDMWFSDHLGHKGFVASM RGFDDPDHEDVILSNFENTIGDDVLTGLDSSGAHRAERLGLIAERLGGVVKH LYPGNHDSCHPMYHAYKRELFLEVDSVOAFOFKWDDVDVLSHPRGQDHGM ESRFDDLRVPLLIHGLHLSQFPMTGQDVGVGAEWGLKPAPELVLQLKWLSELS E Ki"

830..2722

/gene="lcoP"

/locus\_tag="cg2563"

830..2722

/gene="lcoP"

/locus\_tag="cg2563"

/codon\_start=1

/transl\_table=11

/product="ECTOINE BETAINE TRANSPORTER"

/protein\_id="CAF20998.1"

/db\_xref="GI:41326516"

/translation="WSTNSGNLPSQSPSEPHPHDTHPLGVLPGISVDAQRNKFGL DKTVFQTAALIFAFIANGISDPSSVSTMSFWMNTGMLNFVMLIGIGTMYL IAFSRYKLGTDDEPFESFMTAMFGAGIVGIFPFPGSEPLHYLSPPPHYT EGSTPESLHOALQDHFHGLSAGLALVAGALAYSRYRGRVTLISFTFSLFGEK TEGIAGRLLDMALITATPGTATGLSAIQVCGQVQIIISGASEITNNILIAIILIT IGFIISSVCSVKGRILSNLISITGLVLFVFTIGTPIFLNLIIPSVLYGSEFL SWAGLSWGEETIFQAGTAPYAWMTIATWTFVGMFIARISRGTLREFAITWAI PSFIIILATIFGGTATMNRNVDGSSKEQVLFDMFNLSLSTPFIILFVL AVFFTSADSSVMGTMSQGNPAPNKLIVVFWGLCMGIAVVMLLTGESALTGQL NUTIIAIPALVILVMAIATKIDLSSTDPAAIRORYAKAASNAVVRGLIEHGDDEL STEPABEGRGACATFDSADTADITDWTQRTDEBGNVDVDYDFTTGKADGWTPSTEEGE VDAKD"

complement(3014..3295)

/locus\_tag="cg2564"

complement(3014..3295)

/locus\_tag="cg2564"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein\_id="CAF20999.1"

/db\_xref="GI:41326517"

/translation="MATTASKISTIRPAQDALMSVREDLHARFDGLVDPVQVDAILD HVASNRRAKITVFSKIPIAREATAALQIAGNVNADLLDFIALNRGMAA"

3674..4672

/locus\_tag="cg2565"

3674..4672

/locus\_tag="cg2565"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein\_id="CAF21000.1"

/db\_xref="GI:41326518"

/translation="MNDPETHIDRISKEQDPAARSRVQFIVETVRALPNLTTKQAS LAIOLLDAVQLADAAGTKGASTSNASSLPDTFDALTSIGKLDVRSDESWRSFGFQF SETAHPMLIAPEIEIFQHTDVPFGSDAVADPOENQDMWRRLGVTRENLIYKE FSGGQARAVEMLGNWKIGVVVSNTESRLGSLTRVYETTPGHPVPLMSEKNCWY SIRVSTIGENQVPEIVRCLGEIFCGYLFQMWLKEPVRAGKLRIQESAAAYIAWARL DLSPTGTNTWTNSYISTRPLSPAFRWDVLEASHQLENLLRGDTGPTVATQSAAG"

4836..6431

/locus\_tag="cg2566"

4836..6431

/locus\_tag="cg2566"

/codon\_start=1

/transl\_table=11

/product="putative secreted protein"

/protein\_id="CAF21001.1"

/db\_xref="GI:41326519"

/translation="MKKNLLVAPLTASLVPCNLAVANAWEVREASPVVINEVESNS DPGDWVELANTNNNSIDISGSLVDDKEDLENALVLPETGEIISGGFFVIYDSAD YVPTNTPFGQGEYFGLKDDTVTLNRAEYVATYWKDLGHEAENTYGRIPDMTGDF ANTGVPTGAKNVAEGSEBEGVANAQLPFHNVEITPIHLGGFTGEDMSGVDFDA NSTANIANDIKTIVSLAHDIANNITKLTGEWTCYGGGEPDAGEITVAANGDIYL STERNADKNVSRPILSFATPTGKTGVQNAQVOWDLSEFVGDIOPNGLEIAIOLED NIFVGVETGDIVVDSLADQPLVQRYESSFDGWSLSDYNAATKQLSVUCDEACDG LSEILWEPGRKLYSQDIDKIYERPANLGNWAEFGTYISELCKENGVSVTSYILWAD DAATNEGTSLSNAQVINGDCGVNIPEGSSDSSDFATGSIAGAFATVAVLVGIA GALGFFQIILAAFPALQQVIRF"

6556..6942

/locus\_tag="cg2567"

6556..6942

/locus\_tag="cg2567"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein\_id="CAF21002.1"

/db\_xref="GI:41326520"

/translation="MPSFFPRRRKSEPDLLALPAQIRDVHADVLIDIFMAESLVINI DEKSAKILIDAAHHIPTRFTGPNAAPLSVIPIEDPRSRPTLHPDHGWMNPLSPFVVD ELLGGGLKIGTELESTNIAFTVDAS"

complement(7072..8385)

/gene="dctM"

/locus\_tag="cg2568"

complement(7072..8385)

/gene="dctM"

/locus\_tag="cg2568"

/codon\_start=1

/transl\_table=11

/product="PUTATIVE C4-DICARBOXYLATE TRANSPORT SYSTEM (PERMEASE LARGE PROTEIN)"

/protein\_id="CAF21003.1"

/db\_xref="GI:41326521"

/translation="MLSPAVAALIVIGIVVLIASVFAVAIGLPSLFAAMAVLGP ENAQAQORMFTGNTSFTLLAIPEFVLGILMNGSGIATRLIDAKVLVGRMPASMA NTNIAANGLFAGVSGAANAASAVGTWMPKMGKGYRAYAANAANVASAPAGMLIPP SNTFIVLSVTSIAALFMAGVGGGLLWILACVIVGTWLARKENYKREIQIHTPKQS LVLWRLAPLSLMIVIVGGILLGWFTPTESAIAIVVYLCVLGFIYRTIKVGLDAIL LKATRTTIVMLLIAVSALSWMAFAKIPQMSDALLSVDSKVLILIMFILLILI GTVMDPTAIIIFVPIFVTVTELGVDPVHFAGVVMNLSVGVTTPVGNVLFVGSQV AGLRVTIRRLMPYLIATIVALFVVVFVFPQISIWLPPTMGLMGG"

complement(8387..8911)

/gene="dctX"

CDS	/locus_tag="cg2569" complement(8387..8911) /gene="dctX" /locus_tag="cg2569" /codon_start=1 /transl_table=11 /product="PUTATIVE C4-DICARBOXYLATE TRANSPORT SYSTEM (PERMEASE SMALL PROTEIN)" /protein_id="CAF21004.1" /db_xref="GI:41326522" /translation="MNFKSIVCVTAQVFSROVLHSPSTWSELSKLLFWMLSPAGS ALFLGRIHVAIDFIARKLPVSAQRVLQVLIIVFALLGIMVGGIILAASTAMNQ LFLPLTLGLWYVIVFIATGFIATFALFIIDLIIEVATGKERFPYPLVDSERPRDLDELA QSAIDSASAEGRN" complement(8908..9657) /gene="dctp" /locus_tag="cg2570" complement(8908..9657) /gene="dctp" /locus_tag="cg2570" /codon_start=1 /transl_table=11 /product="PUTATIVE C4-DICARBOXYLATE-BINDING PROTEIN" /protein_id="CAF21005.1" /db_xref="GI:41326523" /translation="MMGSVEKKYLRSLFDWFTTPEVGVIGSIHOMSVIRDDLMPEPL FOSLAEKDNISVIGFTQGRNLYTDDAPVKRPADLAGKKIRVOESAMHRLMELGG SATPLTYGVVYTAQMSGLVDGAENNEISVYTONHEFVARYNSNHLVGLDYVWRHD LLDMSBPREFPLEWDAMTEHTDLNWTETDAVIEKAKAGAEFEVDAQAFTDAL APIKDFLTSEFQRELYEAVRAADTSGAAS" complement(9722..11569) /gene="lepA" /locus_tag="cg2571" complement(9722..11569) /gene="lepA" /locus_tag="cg2571" /codon_start=1			
gene				
CDS				
Query Match	99.2%; Score 1515; DB 1; Length 349136;			
Best Local Similarity	99.9%; Pred. No. 0;			
Matches 1526; Conservative	0; Mismatches 0; Indels 1; Gaps 1;			
QY	1	CTCATGCGAATCTGCGCGTTCGGTTCCTGTCAGTGTGGTTCACCGCAACCAAG	60	
Db	340135	CTCATGCGAATCTGCGCGTTCGGTTCCTGCCAGTGTGGTTCACCGCAACCAAG	340076	
QY	61	CGTTTCGGGGCAATCAGTTCCTGGGGCGCGCTATTTGGTATGCGCATGGTTCCTGG	120	
Db	340075	CGTTTCGGGGCAATCAGTTCCTGGGGCGCG--GTATTGGTATGCGCATGGTTCCTGG	340017	
QY	121	CTTGTGTAAACGGCTACGACGTGGCGCCACCATGGCTGGGGGCAAAATGCCAATGTGGTC	180	
Db	340016	CTTGTGTAAACGGCTACGACGTGGCGCCACCATGGCTGGGGGCAAAATGCCAATGTGGTC	339957	
QY	181	CTGTGTTGTTAGATGTTGCCAACCGGTTACAGGGCAACCGTCTCTCTGCTGGT	240	
Db	339956	CTGTGTTGTTAGATGTTGCCAACCGGTTACAGGGCAACCGTCTCTCTGCTGGT	339897	
QY	241	GTTTCTTGGATCTTGGCAACGATCAGAAAGTTCCTGCGCAACGCACTCAAGGGCACTGC	300	
Db	339896	GTTTCTTGGATCTTGGCAACGATCAGAAAGTTCCTGCGCAACGCACTCAAGGGCACTGC	339837	
QY	301	AGACTTCTCTGATCATTCCAGTGTGCTGTGCTGCACCGGATTCCTTACATTCATCGC	360	
Db	339836	AGACTTCTCTGATCATTCCAGTGTGCTGTGCTGCACCGGATTCCTTACATTCATCGC	339777	
QY	361	CATTGCCCCAGCAATCGCTGGGTGGGCGATGTGCTGGCACACGGCTACAGGGAATTTA	420	
Db	339776	CATTGCCCCAGCAATCGCTGGGTGGGCGATGTGCTGGCACACGGCTACAGGGAATTTA	339717	
QY	421	TGATTTTCGGTGTCCAGTCGGCGGCTCTGCTCTTCGGTCTGGTCTACTACCAATCGTCAT	480	
Db	339716	TGATTTTCGGTGTCCAGTCGGCGGCTCTGCTCTTCGGTCTGGTCTACTACCAATCGTCAT	339657	

QY	481	CACGTGCTGTGACCAACAGTCTCTTCCCGCCAAATTGAGCTGGAGCTCTTTAAACAGGCTGATC	540	
Db	339656	CACGTGCTGTGACCAACAGTCTCTTCCCGCCAAATTGAGCTGGAGCTCTTTAAACAGGCTGATC	339597	
QY	541	CTTCATCTTTGCGCAACGGCATCTATGCTAATATCGCCACAGGTCGCGCATGTTTGGCAGT	600	
Db	339596	CTTCATCTTTGCGCAACGGCATCTATGCTAATATCGCCACAGGTCGCGCATGTTTGGCAGT	339537	
QY	601	GTTCTTCTTGGCGAAAGAGTGAAGAAGCTCAAGGGCTTCAGAGTGCCTTCAGGTGTCTCCGC	660	
Db	339536	GTTCTTCTTGGCGAAAGAGTGAAGAAGCTCAAGGGCTTCAGAGTGCCTTCAGGTGTCTCCGC	339477	
QY	661	TGTTCTTGGTATTACCGAGCTCTGCGATCTTTCGGTGTGAACCTTCGCTCGCTGCGCTGCGCT	720	
Db	339476	TGTTCTTGGTATTACCGAGCTCTGCGATCTTTCGGTGTGAACCTTCGCTCGCTGCGCTGCGCT	339417	
QY	721	CTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTGCACTCTTTTAATATCAA	780	
Db	339416	CTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTGCACTCTTTTAATATCAA	339357	
QY	781	GGCAGTTGCGTTGGGCGCTGCAGGTTTCTTGGTGTGTTTCTATTGATGCTCTCAAGATAT	840	
Db	339356	GGCAGTTGCGTTGGGCGCTGCAGGTTTCTTGGTGTGTTTCTATTGATGCTCTCAAGATAT	339297	
QY	841	GGTCATGTTCTTGGTGTGTCAGTTCGTTACCTTCTTTCATCGCATTCGGCGCAGCATTCG	900	
Db	339296	GGTCATGTTCTTGGTGTGTCAGTTCGTTACCTTCTTTCATCGCATTCGGCGCAGCATTCG	339237	
QY	901	TTATGGCCTTTTACTTGGTTCGGCGCAACGGCAGCATTCGATCCAGATGCCAACCCCTGCTCC	960	
Db	339236	TTATGGCCTTTTACTTGGTTCGGCGCAACGGCAGCATTCGATCCAGATGCCAACCCCTGCTCC	339177	
QY	961	AGTGCCCTGCAGGAACGACCAAGCCGAGCAGCAAGCACCCTGCGAGATTTTCAAACGATTC	1020	
Db	339176	AGTGCCCTGCAGGAACGACCAAGCCGAGCAGCAAGCACCCTGCGAGATTTTCAAACGATTC	339117	
QY	1021	CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGACCTCAGCAGCGTCAGCGATGC	1080	
Db	339116	CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGACCTCAGCAGCGTCAGCGATGC	339057	
QY	1081	CATGTTTGGCAGCGGAAAGCTTGGCTCGGGCGTTCATCGTCCCAACCAAGGGGCGAGTT	1140	
Db	339056	CATGTTTGGCAGCGGAAAGCTTGGCTCGGGCGTTCATCGTCCCAACCAAGGGGCGAGTT	338997	
QY	1141	AGTTTCTCCGGTGAGTGGAAAGATTGTTGTTGGCATTTCCCATCTGGCCATGCTTTCCAGT	1200	
Db	338996	AGTTTCTCCGGTGAGTGGAAAGATTGTTGTTGGCATTTCCCATCTGGCCATGCTTTCCAGT	338937	
QY	1201	TGCGACCAAGGCTGAGGATGTTTCCAAATGTTGATATCTTTGATGCACATTGGTTTCCGACAC	1260	
Db	338936	TGCGACCAAGGCTGAGGATGTTTCCAAATGTTGATATCTTTGATGCACATTGGTTTCCGACAC	338877	
QY	1261	AGTAAACCTCAACGGCAGCAGCTTTAAACCCGCTGGAAGAGAGAGGCGGATGAAGTCAAAGC	1320	
Db	338876	AGTAAACCTCAACGGCAGCAGCTTTAAACCCGCTGGAAGAGAGAGGCGGATGAAGTCAAAGC	338817	
QY	1321	AGGGGAGCTGCTGTGTGAATTCGATATTCATGCGCATTAAGGCTGCAGGTTATGAGGTAAC	1380	
Db	338816	AGGGGAGCTGCTGTGTGAATTCGATATTCATGCGCATTAAGGCTGCAGGTTATGAGGTAAC	338757	
QY	1381	CACGCCGATTGTTGTTTTCGAATTTACAAGAAAACCGGACCTGTAAACACACTACGTTTGGG	1440	
Db	338756	CACGCCGATTGTTGTTTTCGAATTTACAAGAAAACCGGACCTGTAAACACACTACGTTTGGG	338697	
QY	1441	CGAAATTTGAAGCGGAGCCAACTGCTCAACCGTTCGCAAAAGAAAGACCGGTGCCAGCAAC	1500	
Db	338696	CGAAATTTGAAGCGGAGCCAACTGCTCAACCGTTCGCAAAAGAAAGACCGGTGCCAGCAAC	338637	
QY	1501	ACCATTAAGTTGAACCTTTCAGTGTTCG	1527	
Db	338636	ACCATTAAGTTGAACCTTTCAGTGTTCG	338610	



Corynebacterineae; Corynebacteriaceae; Corynebacterium.

## REFERENCE

1  
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.

## TITLE

Novel polynucleotides

## JOURNAL

Patent: EP 1108790-A 2904 20-JUN-2001;

## FEATURES

KYOWA HAKKO KOGYO CO., LTD. (JP)

Location/Qualifiers

1. .1983

/organism="Corynebacterium glutamicum"

/mol\_type="unassigned DNA"

/db\_xref="taxon:1718"

## ORIGIN

Query Match 97.7%; Score 1492; DB 6; Length 1983;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1503; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CTATGGCATCTCGCGGTCGCGTTCCTTGGCGCGCGTATTTGATGCGGATGTTTCAACGCAACCAAG 60  
Db 481 CTATGGCATCTCGCGGTCGCGTTCCTTGGCGCGCGTATTTGATGCGGATGTTTCAACGCAACCAAG 540

QY 61 CGTTTCGGCGCAATGAGTTCCTTGGCGCGCGTATTTGATGCGGATGTTTCAACGCAACCAAG 120  
Db 541 CGTTTCGGCGCAATGAGTTCCTTGGCGCGCGTATTTGATGCGGATGTTTCAACGCAACCAAG 599

QY 121 CTTCGTGAACGGCTACGACGTGGCGCGCACCATGGCTGCGGCGCAATGCAATGTGGTC 180  
Db 600 CTTCGTGAACGGCTACGACGTGGCGCGCACCATGGCTGCGGCGCAATGCAATGTGGTC 659

QY 181 CTCTGTTGGTTAGATGTTGCCAAGCCGTTTACCAGGCGACCGTGTCTCTGTGCTGGT 240  
Db 660 CTCTGTTGGTTAGATGTTGCCAAGCCGTTTACCAGGCGACCGTGTCTCTGTGCTGGT 719

QY 241 GCTTCTCTGATCTGCGAAGATCGAGAGTTCCTGCAAGCGACTCAAGGCGACTGC 300  
Db 720 GCTTCTCTGATCTGCGAAGATCGAGAGTTCCTGCAAGCGACTCAAGGCGACTGC 779

QY 301 AGACTTCTGATCACTCCAGTGTGAGTGTGCTCACCGGATTCCTTACATTCATCGC 360  
Db 780 AGACTTCTGATCACTCCAGTGTGAGTGTGCTCACCGGATTCCTTACATTCATCGC 839

QY 361 CATTGCGCCAGCAATGCGTGGTGGCGGATGCTGTGGCAACGCTGTACAGGCACTTTA 420  
Db 840 CATTGCGCCAGCAATGCGTGGTGGCGGATGCTGTGGCAACGCTGTACAGGCACTTTA 899

QY 421 TGATTTGCGTGTCCAGTCGCGGCTGCTCTTGGTCTGCTTACTCAACATGCTCAT 480  
Db 900 TGATTTGCGTGTCCAGTCGCGGCTGCTCTTGGTCTGCTTACTCAACATGCTCAT 959

QY 481 CACTGTCTGCAACAGTCTCTCCGCCAATGAGCTGGAGCTGTTAACCGAGGTGGATC 540  
Db 960 CACTGTCTGCAACAGTCTCTCCGCCAATGAGCTGGAGCTGTTAACCGAGGTGGATC 1019

QY 541 CTTTCATCTTCGCAACGGCATCTATGCTATATATCGCCAGGTCGGGCACTGTTGCGAGT 600  
Db 1020 CTTTCATCTTCGCAACGGCATCTATGCTATATATCGCCAGGTCGGGCACTGTTGCGAGT 1079

QY 601 GTTCTTCTGCGCAAGAGTGAAGCTCAAGGCGCTTGCAGGTCGTTTCAGGTCCTCCGC 660  
Db 1080 GTTCTTCTGCGCAAGAGTGAAGCTCAAGGCGCTTGCAGGTCGTTTCAGGTCCTCCGC 1139

QY 661 TGTTCTTGGTATTAACGAGCTCTCGATCTCTTGGTGTGAACCTTCGCTGCGCTGGCGGCTT 720  
Db 1140 TGTTCTTGGTATTAACGAGCTCTCGATCTCTTGGTGTGAACCTTCGCTGCGCTGGCGGCTT 1199

QY 721 CTTTCATCGGATCGGTACCGCAGCTATCGGTGGCGCTTGAATGCACTCTTTAATATCAA 780  
Db 1200 CTTTCATCGGATCGGTACCGCAGCTATCGGTGGCGCTTGAATGCACTCTTTAATATCAA 1259

QY 781 GGCAGTTCGCTGGCGCTGCGGTTCTTGGGTTGTTTCTATTGATGCTCCAGATAT 840  
Db 1260 GGCAGTTCGCTGGCGCTGCGGTTCTTGGGTTGTTTCTATTGATGCTCCAGATAT 1319

QY 841 GGTCACTTCTTGTGTGTGTCAGTTGTTACTTCTTTCATCGCATTCGCGCGAGGATTC 900  
Db 1320 GGTCACTTCTTGTGTGTGTCAGTTGTTACTTCTTTCATCGCATTCGCGCGAGGATTC 1379

QY 901 TTATGGCCTTTACTTGTGTCGCGCAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCC 960  
Db 1380 TTATGGCCTTTACTTGTGTCGCGCAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCC 1439

QY 961 AGTGCCTTCAGGCAACGACCAAGCCGAGCAGAGCAACCGCGAGATTTTCAACGATTC 1020  
Db 1440 AGTGCCTTCAGGCAACGACCAAGCCGAGCAGAGCAACCGCGAGATTTTCAACGATTC 1499

QY 1021 CACCATCATCCAGCACCTTTTGAACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080  
Db 1500 CACCATCATCCAGCACCTTTTGAACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1559

QY 1081 CATGTTTTCAGCGGAAAGCTTGGCTCGGCGCTTGGCCATCGTCCCAACCAAGGGCGAGTT 1140  
Db 1560 CATGTTTTCAGCGGAAAGCTTGGCTCGGCGCTTGGCCATCGTCCCAACCAAGGGCGAGTT 1619

QY 1141 AGTTTCTCCGTCAGTGGAGAGATTGTTGGTGGCATTTCCCATCTGGCCATCTTTCGCGAGT 1200  
Db 1620 AGTTTCTCCGTCAGTGGAGAGATTGTTGGTGGCATTTCCCATCTGGCCATCTTTCGCGAGT 1679

QY 1201 TCGCACCAAGCTGAGGATGTTTCCATGTGGATATCTTGTATGCATTCATTCGATTCGACAC 1260  
Db 1680 TCGCACCAAGCTGAGGATGTTTCCATGTGGATATCTTGTATGCATTCATTCGATTCGACAC 1739

QY 1261 AGTAAACCTCAACGGCAGCGCATTTTAAACCGCTGAAAGAACGAGGCGGATGAAGTCAAAGC 1320  
Db 1740 AGTAAACCTCAACGGCAGCGCATTTTAAACCGCTGAAAGAACGAGGCGGATGAAGTCAAAGC 1799

QY 1321 AGGGGAGCTGCTGTGATTCGATATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1380  
Db 1800 AGGGGAGCTGCTGTGATTCGATATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1859

QY 1381 CAGCCGATTTGTTTTCGAATTAAGAAACCGGACCTGTAACACACTTACGTTTGGG 1440  
Db 1860 CAGCCGATTTGTTTTCGAATTAAGAAACCGGACCTGTAACACACTTACGTTTGGG 1919

QY 1441 CGAAATGAAGCGGAGCGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTCGAGCAAC 1500  
Db 1920 CGAAATGAAGCGGAGCGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTCGAGCAAC 1979

QY 1501 ACCA 1504  
Db 1980 ACCA 1983

## RESULT 7

BD165105  
LOCUS Novel polynucleotide. 1983 bp DNA linear PAT 17-JAN-2003  
DEFINITION Novel polynucleotide.  
ACCESSION BD165105  
VERSION BD165105.1 GI:27870917  
KEYWORDS JP 2002191370-A/2904.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1983)  
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.  
TITLE Novel polynucleotide  
JOURNAL Patent: JP 2002191370-A 2904 09-JUL-2002;  
COMMENT KYOWA HAKKO KOGYO CO LTD  
OS Corynebacterium glutamicum  
FN JP 2002191370-A/2904  
PD 09-JUL-2002  
PF 15-DEC-2000 JP 2000405096  
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,  
KEIKO OCHIAI,  
PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO

PI OZAKI  
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC  
C12N1/15,  
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/  
04, C12P13/08,  
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC  
G01N33/566,  
PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, C12N1/21, C12R1/15,  
C12N1/21, C12R1/13, C12N1/21, C12R1/01, C12P13/08, C12R1/15,  
PC C12N15/00,  
PC C12N5/00, C12N15/00  
CC Novel polynucleotide  
PH Key Location/Qualifiers  
FT source 1..1983  
FT /organism='Corynebacterium glutamicum'.  
FEATURES  
source  
1..1983  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

## ORIGIN

Query Match 97.7%; Score 1492; DB 6; Length 1983;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1503; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 CTCATGGCATCTCGCGCTTCGGTCTTCCAGTGTGGTTGTTTCAACCGCAACCAAG 60  
DB 481 CTCATGGCATCTCGCGCTTCGGTCTTCCAGTGTGGTTGTTTCAACCGCAACCAAG 540  
QY 61 CGTTTCGGCGCAATCAGTTCCTGGCGCGCGGTATTGTTATGGCGATGGTTCCTCCGAG 120  
DB 541 CGTTTCGGCGCAATCAGTTCCTGGCGCGG-GTATTGTTATGGCGATGGTTCCTCCGAG 599  
QY 121 CTTGTTGAACCGGTACGACGTGGCCGCCACATGTCGTCGGCGGCAAAATCCCAATGTGTC 180  
DB 600 CTTGTTGAACCGGTACGACGTGGCCGCCACATGTCGTCGGCGGCAAAATCCCAATGTGTC 659  
QY 181 CTTGTTGTTTGTAGTGTTCGCCAGCCGTTTACAGGCGACCGTCTTCTGTGCTGT 240  
DB 660 CTTGTTGTTTGTAGTGTTCGCCAGCCGTTTACAGGCGACCGTCTTCTGTGCTGT 719  
QY 241 GCTTTCTTGGATTCTGGCAACGATCGAAGATTCCTGCAACAAGCGACTCAAGGGCACTGC 300  
DB 720 GCTTTCTTGGATTCTGGCAACGATCGAAGATTCCTGCAACAAGCGACTCAAGGGCACTGC 779  
QY 301 AGACTTCCTGATCACTCCAGTGTGCTGCTTCTGCTCAGCGGATTCCTTACATTCATCCG 360  
DB 780 AGACTTCCTGATCACTCCAGTGTGCTGCTTCTGCTCAGCGGATTCCTTACATTCATCCG 839  
QY 361 CATTTGCCCGCAGCAATCGCTGGTGGCGATGTCGTCGACACGCGTCTACAGGGACTTTA 420  
DB 840 CATTTGCCCGCAGCAATCGCTGGTGGCGATGTCGTCGACACGCGTCTACAGGGACTTTA 899  
QY 421 TGATTTCCGTTGTCCTGCGGCGTCTGCTCTTCCGTTCTGCTTCTGCTTCTGCTTCTGCT 480  
DB 900 TGATTTCCGTTGTCCTGCGGCGTCTGCTCTTCCGTTCTGCTTCTGCTTCTGCTTCTGCT 959  
QY 481 CACTGTCCTGACCACTCTTCCGCGCAATTGAGCTGGAGCTGTTTAAACAGGGTGGATC 540  
DB 960 CACTGTCCTGACCACTCTTCCGCGCAATTGAGCTGGAGCTGTTTAAACAGGGTGGATC 1019  
QY 541 CTTTCATCTTCGCAACCGGATCTATGCTAATATCGCCAGGGTGGCGCATGTTTGGCAGT 600  
DB 1020 CTTTCATCTTCGCAACCGGATCTATGCTAATATCGCCAGGGTGGCGCATGTTTGGCAGT 1079  
QY 601 GTTCTTCTGCGGCAAGTGAAGCTCAAGGCGCTTGCAGGTGCTTCAAGGTCTTCCGC 660  
DB 1080 GTTCTTCTGCGGCAAGTGAAGCTCAAGGCGCTTGCAGGTGCTTCAAGGTCTTCCGC 1139  
QY 661 TGTCTTGTGTTATACGAGCTCGGATCTTCCGTGTGAACCTTCGCTGGCTGGCGGTT 720  
DB 1140 TGTCTTGTGTTATACGAGCTCGGATCTTCCGTGTGAACCTTCGCTGGCTGGCGGTT 1199

QY 721 CTTTCATCGGTATCGGTACCGAGCTATCCGTGGCGCTTTGATTGCACCTCTTTAATATCAA 780  
DB 1200 CTTTCATCGGTATCGGTACCGAGCTATCCGTGGCGCTTTGATTGCACCTCTTTAATATCAA 1259  
QY 781 GGCAGTTCGTTGGCGCTGCAGGTTCTTGGGTGTTTCTTATGATGTCGCAAGATAT 840  
DB 1260 GGCAGTTCGTTGGCGCTGCAGGTTCTTGGGTGTTTCTTATGATGTCGCAAGATAT 1319  
QY 841 GGTCAATGTTCTTGGTGTGTCAGTGTGTACCTTCTTCACTCCGATTCGCGGCGAGGATTCG 900  
DB 1320 GGTCAATGTTCTTGGTGTGTCAGTGTGTACCTTCTTCACTCCGATTCGCGGCGAGGATTCG 1379  
QY 901 TTATGGCCTTTACTTGGTTTCGGCCACAGCGGAGCAATTGATCCAGATGCAACCGCTGCTCC 960  
DB 1380 TTATGGCCTTTACTTGGTTTCGGCCACAGCGGAGCAATTGATCCAGATGCAACCGCTGCTCC 1439  
QY 961 AGTCTGTCGAGAACGACGACAAAGCCGAGCAGACACCGCAGAAATTTTCAACGATTC 1020  
DB 1440 AGTCTGTCGAGAACGACGACAAAGCCGAGCAGACACCGCAGAAATTTTCAACGATTC 1499  
QY 1021 CACCATCATCCAGGCACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080  
DB 1500 CACCATCATCCAGGCACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1559  
QY 1081 CATGTTTCCAGCGGAAAGCTTGGCTCGGGCGTTGCCATTCGTCCTCCAAACCAAGGGCAGTT 1140  
DB 1560 CATGTTTCCAGCGGAAAGCTTGGCTCGGGCGTTGCCATTCGTCCTCCAAACCAAGGGCAGTT 1619  
QY 1141 AGTTTCTCGGTGAGTGGAAGATTTGGTGGCATTCCTATCTGGCCATGCTTTCGCAAT 1200  
DB 1620 AGTTTCTCGGTGAGTGGAAGATTTGGTGGCATTCCTATCTGGCCATGCTTTCGCAAT 1679  
QY 1201 TCGCACCAAGCTGAGGATGGTTCCAAATGCTGATATCTTGATGACATTTGGTTTCGACAC 1260  
DB 1680 TCGCACCAAGCTGAGGATGGTTCCAAATGCTGATATCTTGATGACATTTGGTTTCGACAC 1739  
QY 1261 AGTAAACCTCAACGGCAGCCTTTAAACCCGCTGAAGACGAGGCGATGAAGTCAAAGC 1320  
DB 1740 AGTAAACCTCAACGGCAGCCTTTAAACCCGCTGAAGACGAGGCGATGAAGTCAAAGC 1799  
QY 1321 AGGGAGCTGCTGTGTGTAATTCGATATTCGATCCGATTAAGGCTGCAAGTTATGAGTAAC 1380  
DB 1800 AGGGAGCTGCTGTGTGTAATTCGATATTCGATCCGATTAAGGCTGCAAGTTATGAGTAAC 1859  
QY 1381 CACGCGGATTTGTTTTCGAATTAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440  
DB 1860 CACGCGGATTTGTTTTCGAATTAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1919  
QY 1441 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACGTTCGCAAAAGAAAGAGCGGTGCCAAGC 1500  
DB 1920 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACGTTCGCAAAAGAAAGAGCGGTGCCAAGC 1979  
QY 1501 ACCA 1504  
DB 1980 ACCA 1983

## RESULT 8

BD093238  
LOCUS DNA encoding sucrose PTS enzyme II.  
DEFINITION  
ACCESSION BD093238  
VERSION BD093238.1 GI:22638826  
KEYWORDS WO 0102584-A/1.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 5969)  
AUTHORS Izui, M., Sugimoto, M., Nakamatsu, T. and Kurahashi, O.  
TITLE DNA encoding sucrose PTS enzyme II  
JOURNAL Patent: WO 0102584-A 1 11-JAN-2001.  
AJINOMOTO CO INC,MASAKO IZUI,MASAKAZU SUGIMOTO,TSUYOSHI NAKAMATSU,



Db	721	TTAGTTTCTCCGTCAGTGGAAAGATTGTGGTGGCAATTCCTCCCATCTGGCCATGCTTTCCGCA	780
QY	1199	GTTGCGCACCAAGGCTGAGGATGGTTCCAAATGTGATATCTTTGATGCAATTTGTTTGGAC	1258
Db	781	GTTGCGCACCAAGGCTGAGGATGGTTCCAAATGTGATATCTTTGATGCAATTTGTTTGGAC	840
QY	1259	ACAGTAAACCTCAACGGCAGCAGCTTTAAACCGCTGGAAGACGAGCGGATGAAGTCAAA	1318
Db	841	ACAGTAAACCTCAACGGCAGCAGCTTTAAACCGCTGGAAGACGAGCGGATGAAGTCAAA	900
QY	1319	GCAGGGGAGCTGCTGTGTGAATTGATATATGATGCAATTAAGGCTGCAAGTTATGAGGTA	1378
Db	901	GCAGGGGAGCTGCTGTGTGAATTGATATATGATGCAATTAAGGCTGCAAGTTATGAGGTA	960
QY	1379	ACCAAGCCGATTTGTTTTCGAATTACAGAAAAACCGACCTGTAACACATTCACGTTTG	1438
Db	961	ACCAAGCCGATTTGTTTTCGAATTACAGAAAAACCGACCTGTAACACATTCACGTTTG	1020
QY	1439	GGCGAAATTGAAGCGGGAGCCAACTGCTCAACCTGCGCAAGAAAGACGCGTCCAGCA	1498
Db	1021	GGCGAAATTGAAGCGGGAGCCAACTGCTCAACCTGCGCAAGAAAGACGCGTCCAGCA	1080
QY	1499	ACACCATAGTTGAAACCTTGAGTGTTCG	1527
Db	1081	ACACCATAGTTGAAACCTTGAGTGTTCG	1109
RESULT	10		
AL935252/c		343050 bp	DNA linear BCT 21-FEB-2003
LOCUS			
DEFINITION			Lactobacillus plantarum strain WCFS1 complete genome; segment 1/11.
ACCESSION		AL935252	
VERSION		AL935252.1	GI:28269805
KEYWORDS			
SOURCE			Lactobacillus plantarum WCFS1
ORGANISM			Lactobacillus plantarum WCFS1
			Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
			Lactobacillus.
REFERENCE			1
AUTHORS			Kleerebezen, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O. P., Leer, R., Tarchini, R., Peters, S. A., Sandbrink, H. M., Piers, M. W. B. J., Stiekema, W., Lankhorst, R. M. K., Bron, P. A., Hoffer, S. M., Groot, M. N. N., Kerkhoven, R., de Vries, M., Ursing, B., de Vos, W. M. and Siezen, R. J.
TITLE			Complete genome sequence of Lactobacillus plantarum WCFS1
JOURNAL			Proc. Natl. Acad. Sci. U.S.A. 100 (4), 1990-1995 (2003)
MEDLINE			22480296
PUBMED			12566566
REFERENCE			2 (bases 1 to 343050)
AUTHORS			Kleerebezen, M. and Siezen, R. J.
TITLE			Direct Submission
JOURNAL			Submitted (14-OCT-2002) Wageningen Centre for Food Sciences, P.O. Box 557, 6700 AN Wageningen, The Netherlands
FEATURES			Location/Qualifiers
source			1. .343050
			/organism="Lactobacillus plantarum WCFS1"
			/mol_type="genomic DNA"
			/strain="WCFS1"
			/db_xref="taxon:220668"
			/notes="single colony isolate of strain NCIMB8826"
gene			1. .1368
			/gene="dnaA"
			/locus_tag="lp_0001"
CDS			1. .1368
			/gene="dnaA"
			/locus_tag="lp_0001"
			/codon_table=1
			/transl_table=11
			/product="chromosomal replication initiation protein DnaA"
			/protein_id="CAD62703.1"
			/db_xref="GI:28269806"
			/db_xref="TrEMBL:Q890K8"
			/translation="MLERNDLWNLKFSNQDLSKITFTQFVEPAKPLQLDKAQMTEI



```
gene      10837..11073
/locus_tag="lp_0011"
/locus_tag="lp_0011"
/locus_tag="lp_0011"
/locus_tag="lp_0011"
/codon_start=1
/transl_table=11
/product="ribosomal protein S18"
/protein_id="CAD62711.1"

Query Match      20.2%; Score 308.4; DB 1; Length 343050;
Best Local Similarity 60.3%; Pred. No. 1.4e-71;
Matches 552; Conservative 0; Mismatches 351; Indels 13; Gaps 2;

QY 4 ATGGCAATCGCCCGCTTGGCTTTCGACAGTGGTGGTTTTCACGCAACCAAGCGT 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 TTCCGGCGCAATGAGTTCTCGGGCGCGCTATTTGGTATGGGATGGTGTTCGCCGAGCTT 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 124 GGTGAACGCTACGAGTGGCGGCCACCATGGCTCGGGCGCAATGCCAATGTGTCCCT 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166832 AGTTAACGCTATAGGTTGCAACGACCATGGCTGCTGGCAAGATGTTTACTTGAACGT 166773
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 184 GTTGGTTTAGATGTTGCCAAGCCGGTTACAGGCGACCGTCTTCCTGCTGCTGGTGT 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166772 CTTTGGTTTACAGTTGCAACAGCAGGCTATCAAGGCCAAGTGCCTCTGTTTGGCCGT 166713
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 244 TTCTTGGATTTGGCAACGATCGAAGTCTCTGCAAGGACACTCAAGGGCACTGCAGA 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166712 TCGTTATATTTGGCTACTTTGGAGAGTCTTCCAAACATATCAAGGGTGCATTGA 166653
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 304 TTTCCTGATCACTCCAGTCTGAGCTGTGTCTGCTCAACCGGATTCCTTACATTCATGCCAT 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166652 CTTTACGTTTACGCCAATGTTTGAATTTGCTATCTACCTGCTTCTTAACTTTCACATCGT 166593
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 364 TGGCCGCAAGTCGCTGGTGGCGATGCTGCGACACCGTCTACAGGACTTTATGA 423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166592 CGGTCTCTGTTTACGAACCGTGAGCGACGATTAACATAACGGTTAGTAGGCTTTATATA 166533
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 424 TTTCGCTGTCTCAGTCGGCGGCTGCTCTCTGCTGCTGCTCTACTACCAATCGTATCAC 483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166532 CAGGCTGTTGGATGGCATGGGATCTTTGGCCCTTCTACTCGGCATTTGCTATTAC 166473
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 484 TGGTCTGCACAGTCTCTCCGCCAATTCAGCTGGAGCTGTTT-----AACCA 531
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166472 TGGGCTCCACCAAACTCTCCGGCTATTGAAACCCAGTTATTGGCTAACGTCGCCAAAAC 166413
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 532 GGTGGATCTTTCATCTTTCGACCGCATCTATGGCTATATATGCCAGGCTGCGCATG 591
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166412 GGGAGGTTCTTTTATCTTCCAGTCGCCCTCGATGGCCACATTTGGCAAGGTGCGCCAC 166353
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 592 TTTGGCAGTGTCTTCTCTCGGCAAGAGTCAAGAGCTCAAGGGCTTGCAGGTGCTTCAGG 651
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166352 TTTAGCTATTTCTTTGGCCTAAGAGCCAAAGCAAAAGCCCTGACTTCTTCAGCCGG 166293
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 652 TGTCTCCGCTGTTCTTGGTATTACGAGCGCTCGATCTTTGGTGGTGAACCTTCGCCCTGG 711
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166292 GGTCTCAGCGTTACTCGGAATACGGAACCTGCTATTTTCGGGGTCAACCTCAAGATGA 166233
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 712 CTGCGGCTTCTTCATCGGTATCGGTACCGCATCTCGGTGGCGCTTGTGATTCATCTT 771
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166232 ATTTCATTTGCTTTTCAGAGCATTCCTCAGGAATGCTCTAGCTTCTTAGACTATT 166173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 772 TAATATCAAGGAGTGTGCTTGGGCGCTCGAGGTTTCTTGGGTGTTTCTTATTGATGC 831
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166172 CCAGGTTCTGCTCGTGAGATGGTCCCGCATCAGTCATCGGGTTCATTTCGATTGCTTC 166113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 832 TCCAGATATGGTCATGTTCTTGGTGTGTCAGTTGTACCTTCTTTCATCGCATTCGGCG 891
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
/gene="rafp"
3909..3914
/gene="rafp"
/note="putative"
3932..3937
/gene="rafp"
/note="putative"
3993..3918
/gene="rafp"
/function="raffinose transport protein"
/note="putative"
/codon_start=1
/transl_table=11
/product="permease"
/protein_id="AAA25563.1"
/db_xref="GI:475108"
/translation="MQEENHYKVGGRLLYFGKAKNDAPYSILSGYLIFITSHLPD
TGNKLDNRMSVLTLLIMVRLVEIFDIPFGNAIDRTKNSPGHFRPWVVGTVSS
IILLFTNLGLYAKNMIYLVFAILYITMDIFYSFKDVGFWMLPSLTTDSRRE
KATFARLSTIGGLVGLVMPAVIFFSAKATSTGDNKGFIFALIICLIASIWG
VLGTREVDSDIRKNDQTVGMVEIFKALAKNDOLLAAALYLFYGVGINILGLSVY
YFTYMGKPKSPSILSINIFLGLIATSLFPVLKSPKSGVAGCLVFMGLGIAFT
IAGSNLWLLAATMGFPQOMVPLVLMVITDSVYQGLKLRHDSIALSVRPLID
KFGCALSNVGGQIALISGMTTGATASSITAAGQLHFKLTMFAFPAIMLLIIAIFGSK
LITTEKKAIEIVAELEIRWTRTKFNTDQVAKVYVLTSLDPAIAGQVPLAQVNDP
TPAAGTIGDGFAPKPSDGRILAPFDATVRQVTRHAGVLGNGIVLIIHIGLTVK
LRGTGISVVEGQHVQQGDELLEFWDPTIKQAGLDDTIVMTVTNSTEFTMDWLKP
GQAVKATNIIQLHTKA"
5976..8177
/gene="aga"
5976..8177
/gene="aga"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="alpha-galactosidase"
/protein_id="AAA25564.1"
/db_xref="GI:475109"
/translation="MSLIIVDQANRVFHLHNOTLSYI PAVEQGGTLSHLYFGGVNDHY
HGLRYRVRDGRFSGNLPGSTDRTPSRDTPKEYTAGSMDYHLPAAIVRHDTGANAL
VLVYQYRLIAGKPKLSGLPAFVETEAETLTIIVLQVQSVQFEDLQTIYDRPV
VTRSVQVQNGDRAVNLKVASQIDTDRQFETITLPGAHANERPERGSINYQIT
FQSLRTSSHQWNPFLALVDHTTTEFSGDAYGNLYSGNHAPELEKQDQLHLMVG
INSYNFQWLKAGATQTEPELVMTVYTKNGIAMSQYHHLIRVVRVSEFKQNERPIV
VNNREATFDFNEAKLPIVDEAKOLGEMFVLDDGFGHRDDNSSLGSDGWDVDEKPE
POLNHFVYVHEQGLKFLWLEPEMI SYDSLKYQHPDYLMOVGRSPSRNOXIL
DLGROAVRNIIIDOLLKSKOLDYIKDMRHLSDIYSVALPPERQEVHRYVILG
YELLERLTAIPHILFEGSCGGGRFDAGMATYMPQI WASONTIDAVARLTIQYGTSL
AFISIALTHVSVEFPNQQTSRSTMSRASAASVGLYELDLTQLSSADKQIVQKQV
VOYKRLPIQFGEFVQKSPITTSNOAAMFVSPQDEAIVMVFNLTSYAQPSLATKTK
LVGLNPKLNYQNIATKAI FGGDELMQLGFYDPVYQDYTKVYHFKAVTEN"
8265..8289
/standard_name="terminator 3"
/note="putative"
9403..9426
/note="DR2; putative"
/rpt_type=direct
10396..10442
/note="DR1; putative"
/rpt_type=direct
10397..10443
/note="IR2; putative"
/rpt_type=inverted
10681..10716
/standard_name="terminator 2"
/note="putative"
complement(10727..11593)
/gene="scrk"
complement(10727..11593)
/gene="scrk"
/function="fructose 6-phosphotransferase"
/note="putative"
/codon_start=1
```

```
/transl_table=11
/product="fructokinase"
/protein_id="AAA25565.1"
/db_xref="GI:475110"
/translation="MLLGAIEAGCTKFKVCATGAENGQVSDRI SIPTTTPVETMTAVDD
YFTTHPDAIGISFGPIGVNPHDPKYGIITTPKQWGDFFDLGHLKSNFNIPLYWT
TDVNEAAYGESMIGIAKDVNSIYMTIGTGAGAVI SQNHI FNGRTHTEIGHMELNEL
PGDFFKSNCPYHDIICLEGAAAGPAVKRTGKAGKDI PVDDPWPPIITDIYAQACVNLT
VAFAPKIIILNGVGNQRFQPMIRKFAAYLNGYEVEVPLDDIIVPAGLGNNSGIAG
GLLLAQAALKNA"
complement(11673..13832)
/gene="agas"
complement(11673..13832)
/note="agas"
/note="putative"
/codon_start=1
/transl_table=11
/product="alpha-galactosidase"
/protein_id="AAA25566.1"
/db_xref="GI:475111"
/translation="MPVEYDPKTCGLINLHNDQISYVIOILAHRYPHRYFGRVFSKQP
YFAPESGSHAFANDPTEFPYVSVSLPSEYSTIGSDYRQPAIVTKDNQLLPILIE
YTFVNDQPTINSRQLPPTVSKHTPTVITVILHTDAVKIQMDLNTIIFENQLILRS
TLRHGTTLQVLTALSSAQLDPTDQYGTALTSHTAHEANPSFNLHPGLQTVRSI
RTSGQHOPFMALAEPTTELATGIGCALAWSGNPDSVVDVQYHSLRTILGLPD
TFEMLKPNSSFCOTPEAVLTWTFNGMSOVPHDFSOLMPSQMTIPSVLNTWEIT
FAVESKVOHLIEHAQLGLQMLVLDGDFVNRNGENGLGDFWFDPIKFPNGNLPIA
QQAHHRMKFLWVEFEIMTINSQLYQQHPDWLVYQVTRPTITARHQLVDLSQAAR
DHLITLILNQNLQDLKLMNRLHTQVSTHLPAAQQGELYHRYVCGLYDILTRL
KRACPKLIENGAGGRDFGMLPYTNQITWISDLDPVDRATIENGFSYLPPIRFS
NHITSPNAQNGRI TPETRLQACIGLGLNPKOLAPSEQOILARGALIKYQOLKS
TPIKAHYRLTPTRHVAVLI VTADKKQALCVLNGLSRVKTHQHPILPHYLDLALY
SDSSGNRYTHQNLTMGIPKPTNADPTSQLIYLQN"
complement(13852..15989)
/gene="scrA"
complement(13852..15807)
/gene="scrA"
/function="enzyme II of the PTS system, sucrose specific"
/note="enzyme IIabc"
/codon_start=1
/evidence=experimental
/transl_table=11
/protein_id="AAA25567.1"
/db_xref="GI:475112"
/translation="MNHQEVADRVLNAIGKNNIQAHAACATRLILVIKDSKIDQAL
DDADVQDFETNGQIIIGEGVDKVDALIVKTLKEVTPDDIKAAAGQKNKP
LMDFLKVSDFIPFIPALVAGLLMALNNVLTAEHLFWAKSVVEYVYGLKGAEMIN
AMASAPFTPLILGFSATKRGPNPLGATGMIMVLPSLVNGSVATVMAAGKMY
MNVFGLHVAQAGYQGVLPVGLVAFILATLEKFFHKKHKGADFTTFMFAIVITGFL
TPTIYVPLTVSDALTNGLVNLTGWMIGLGLYSAIVITCLHOTEPAITQL
LANVAKTGSFTFPVASMNIQGAATLAFATKSQKQKALTSAGVSAALGITPA
IFGVNLKMPFPVFAIASIASIAGLFLHLSVAMSPASVIGFISASKSPAFMFLS
AVISVAFIPFTPIYAKRTLDGDRDQVKSAPTSTVTVNVDIISAPVTCASELSQ
NDQVFAEIMGKAAIVPSSDQVAPADGVIITVYDSHAYGKITTAGAEIILHLGLD
TVNLNGEHTTVNQKQDTHQGLGTGFDIAALKAAVNDPTVWMLIVTNTANVBERL
KVTNVQAGEOLVALTAPAASSVAATTV"
complement(15872..15877)
-10_signal
/gene="scrA"
complement(15894..15989)
-35_signal
/gene="scrA"
15993..17568
/gene="scrB"
15993..15998
-35_signal
Query Match 18.8%; Score 287.6; DB 1; Length 21838;
Best Local Similarity 58.8%; Pred. No. 4.3e-66;
Matches 539; Conservative 0; Mismatches 364; Indels 13; Gaps 2;
QY 4 ATGGCATCTGCCCGCTTCGCGTTCTTCCAGHGTGGTGGTTTCCCGCAACCAAGCGT 63
|||||
Db 15324 ATGGCTAGTGGCGCGCTTACTTTCTTCCCAATCTCTTAGGATTTTCAGCAAGAGCGC 15265
|||||
QY 64 TTCGGCGGCAATGAGTTCCTCGGGCGCGCGTATTGGTATGGCGATGGTGTTCGCCGAGCTT 123
```

Db	15264	TTGGTGGCAACCCCTATCTGGTGTCTACG-ATGGCATGATCATGGTCTTACCATCACT	15206	AUTHORS	Leenhouts,K.K.J., Bolhuis,A.A., Kok,J.J. and Venema,G.G.
Qy	124	GGTGAACGGCTACGACGTGGCGCCACCATGGCTGGCGGCGAAATGCCAATGTGGTCCCT	183	JOURNAL	The sucrose and raffinose operons of <i>Pediococcus pentosaceus</i> PPE1.0
Db	15205	AGTTAACGGCTATAGCGTTGCAACACCATGGCAGCGCGAAGATGGTTACTGGAACT	15146	REFERENCE	2 (bases 1 to 21839)
Qy	184	GTTTGGTTAGATGTTGCCAAGCCGGTTACACAGGACCGTGTCTCGTGTGGTGGT	243	AUTHORS	Leenhouts,K.K.
Db	15145	CTTTGGGTACACGTTGACAGCCGGCTATCAAGCCCAAGTCTCCAGACTTTAGGCGT	15086	JOURNAL	Direct Submission
Qy	244	TTCTTGGATTCTGGCAACGATCGAAGAGTTCTCTGCACAAAGCACTCAAGGGCACTGCAGA	303	COMMENT	Submitted (27-APR-1994) Leenhouts K. K., Institute for Biological
Db	15085	CGCCTTCATTAGTACCCCTTGAAAAATCTTCCATAAACACATTAAGGGCGATTGA	15026	FEATURES	Sciences, Department of Genetics, Kercklaan 30, Haren, The
Qy	304	CTTCTGATCACTCCAGTGTGCTGACGTTGTGCTGCTACCGGATCTCTTCAATCAATGCCAT	363	source	Netherlands, 9751 NN
Db	15025	CTTACGTTTACCCCGATGTTTGCATTTGATTAAGTCTGTTTCCAACTTTTACAATCGT	14966	repeat_unit	On May 26, 1994 this sequence version replaced gi:475962.
Qy	364	TGSCCCAGCAATCGCTGGTGGCGATGCTGTGGCACACGCTCTACAGGCACTTTATGA	423	repeat_unit	Location/Qualifiers
Db	14965	TGGTCCCGTCTTACGAACGGTGAGCAATGATTAACGCTTTAGAGCTTATATAA	14906	repeat_unit	1. /organism="Pediococcus pentosaceus"
Qy	424	TTTCCGTGTCCAGTCCGGGCTCTGCTCTCTCGGTCTGCTTACTCACCAGTCTCATCAC	483	repeat_unit	/mol_type="genomic DNA"
Db	14905	CAGCACCGGCTGATTTGGTATGGAATCTTTGTTTATATATCTGCAATTTGTTATTAC	14846	repeat_unit	/strain="PPE1.0"
Qy	484	TGGTGTGACACGATCTTCCCGCCCAATGAGCTGAGCTGTT-----TAACCA	531	terminator	/db_xref="taxon:1255"
Db	14845	TGGTCTCCATCAAACTTTCCAGCAATCGAAACCCAGTTGTTGGCCAAATGTTGCCAAAAC	14786	gene	487..510
Qy	532	GGGTGATCTTCATCTTCGCAACGCACTATGCTATATATGCTATATCGCCAGGGTGGCGATG	591	CDS	/note="DR2"
Db	14785	TGGTGGTTCCTTATCTTCCAGTCCCTCGATGGCCAAACATTTGGCAAGGTCGCGCCAC	14726		/citation=[1]
Qy	592	TTTGGCAGTGTCTTCTCGCGAAGAGTCAAAAGCTCAAGGCCCTTGCAGGTGCTTCAGG	651		/rpt_type=DIRECT
Db	14725	TTTAGTATTTCCTTGGCCACTAAGAGCCAAAGCAAAAGCCCTGACTCTTCAGCCGG	14666		513..561
Qy	652	GTCTCCCGTGTCTTGGTATACGAGCTCGGATCTTCGGTGTGAACTTCGCTCGCTGG	711		/note="DR1"
Db	14665	GGTCTCAGCGTTACTCGGAATACGGAACCTCTATTTTCGGGGTCAACCTCAAGATGAA	14606		/citation=[1]
Qy	712	CTGGCGGTTCTTCATCGGTATCGGTACCGCAGTATCGGTGGCGCTTCGATTCGACTCTT	771		/rpt_type=DIRECT
Db	14605	ATTTCATTGCTTTTGGACGCAATGCCCTCAGGAATGCTTCAGCTTCTTAGGACTATT	14546		2724..2760
Qy	772	TAATATCAAGCAGTTCGTTGGGCGCTGACGTTTCTTGGGTGTTGTTTCTATTGATGC	831		/standard_name="terminator 4"
Db	14545	CCACGTTCTGTCGTCGGCATGGTCCCGCATCAGTCATCGGATTCATTTTCGATTGCTTC	14486		/citation=[1]
Qy	832	TCCAGATATGGTCATGCTTCTTGGTGTGTCAGTTGTTACTCTTTCATCGCATTCGGCGC	891		/product="regulator"
Db	14485	AAAGTCGATTCGACGATTCATGCTCAGTGTGTCATCTCGTTCGTGCTGCAATTATTC	14426		/protein_id="CAA83663.1"
Qy	892	AGCGATTGCTTATGGC 907			/db_xref="GI:475963"
Db	14425	AACTTTTATCTAATGCC 14410			/db_xref="GOA:P43465"
RESULT 12					/translation="MNGEYKTLANKSFESNVLFQGEACLPNYTKGNVRDSVYHI
PPURFOP/c					TOEGKGTAAAHNPATVLKAGDIFILPKGTFCFOADNDOPWKYFWIGFSAGIRIEAM
LOCUS					LSGSLAQKCVLRQYONGHIVADLSLYKVLHNPNSLINDVLGSLIYRLFYDLLRWY
DEFINITION					PADATNKVSTEQENLAVSYLOENYSTGCTIMDLCHLNLRSVLYTLFKTHANTSP
ACCESSION					QKLLPKLRLEDKARLSTNNNSVQSIANNVGVKDSFTFSKAFKRYSGASPSYRKSIG
VERSION					1"
KEYWORDS					complement(3837..3842)
SOURCE					/gene="rafr"
ORGANISM					/citation=[1]
REFERENCE					complement(3859..3864)
					/gene="rafr"
					/citation=[1]
					3909..5918
					/gene="rafr"
					/citation=[1]
					3909..3914
					/gene="rafr"
					/citation=[1]
					3932..3937
					/gene="rafr"
					/citation=[1]
					3993..5918
					/gene="rafr"
					/function="raffinose transport protein"
					/standard_name="permease"
					/citation=[1]
					/codon_start=1
					/transl_table=11
					/product="permease"

```

/protein_id="CAA83664.1"
/db_xref="GI:475964"
/db_xref="GOA:P43466"
/db_xref="Swiss-Prot:P43466"
/translation="MQEERNYKVGGRLLYGFCAKNDAPYSILSGVLIIFITSHLFD
TGNKLDNRVSLVTLIMVIRIVELFDIFIGNADIRTSNCPHFRPWVVGTVSS
IILLFLTLNGLGGLVAMLYLVFAILYITMDIFSKDVGFWMLPSLTDSRRE
KTATFARLSTIGGGVGLVAVPAVIFPSAKATSGDNRGWFIPTALICILISAWG
VGLTREVSDSIDRKNQDVGVMIEIPKALAKNDQLLAAVLAVFYGVGINLGLSVY
YFTVINGKPKSFSLISINILFLGLIATSLPVLSSKFSKGVFAGCLVFMGLGIAFT
IAGSNLWLVLLAATMGFFQOQVFLVLMVITDSVEYQKLGRHDSIALSVRPLID
KFGGLSNGVVGQIAGISMTTGATASSITAAQQLHFKLTFAPFALMLLIATGIFSK
QPLTFEKGABIVAEIERWTRKFDNTDQAEKVTSLDLATPIAGQVPLAQVNDP
TFAAGTLGSPDGRILAPDAROVFTTRHAGLVGNGVILVLIHIGLTVK
LRGTGFIISVVEGQHQQGDDELLEFWDPTFKQAGLDDTVIMTVYNSTFTMDWLVP
GOAVKATDNLQLHTKA"
5976..8177
/gene="agar"
5976..8177
/gene="agar"
/standard_name="alpha-galactosidase"
/citation=[1]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="alpha-galactosidase"
/protein_id="CAA83665.1"
/db_xref="GI:475965"
/db_xref="GOA:P43467"
/db_xref="Swiss-Prot:P43467"
/translation="MSLITVDQANRVFHLNQLTSYFVAVEQGTLSHLVFGHVDHY
HGLRYPRVDRGFSNLPSTDRTEFRDTPKEYSTAGBMDYHLPAIVRHDTGANAL
VLVYQGVIEAGKPLSGLPAAFEVDETEAETLTVLQVQSVDFDLOYTVIRDRPV
VTRSVQCNQGHAVNLEKVASQIDFTDRQETITLPGAHANRPEGRSINYGIQT
EGLRSTSHOMPFIALVDHTTETSGDAYGNLVYSGNHAFELEKQDQLHLMVG
INSYNFQKAGATQTEPVLVMVYTNKGLNAMSQAYHLHILRERVRSFBNQERPIV
VNWAEATDFENAKLPIDEAKQLGIEMFLDDGFGHRRDSDSLGDWQVDRKPF
POLGNHFVYVHQHGLKGLIPEMISYDSKLYQOHPDYLMOVGRKSPSPRNOYL
DLGQANVRNIPDQLDLKSKQIDYIKWDMNRHLSDIYSVALPPERQGVVHRVYLG
LYELLERLTAYPHILFEGCSGGGRFDAGMAYMQIWAASDNDAVARLITQYGTSL
APYLSATHAVSVSNQQTGREGSTRSAVAGSVGLGELDTQLSSADKQIVQKQV
VQTKRLPLIQGEFRLKSPITSNOAAWVFSPQDEAIVMVFNLTSYAQPSLRTK
LVGLNPLKYNQNIATKAIFGGDELMQLGFDVYVQDYTKVYHFAVTEN"
8265..8289
/standard_name="terminator 3"
/citation=[1]
9403..9426
/notes="DR2"
/citation=[1]
/rpt_type=DIRECT
/complement(9519..10424)
/gene="iso-IS30"
/complement(9519..10424)
/gene="iso-IS30"
/function="IS-30 (E.coli) like IS element"
/standard_name="IS-element"
/citation=[1]
/codon_start=1
/transl_table=11
/product="insertion element"
/protein_id="CAA83666.1"
/db_xref="GI:475966"
/db_xref="GOA:Q52261"
/db_xref="TrEMBL:Q52261"
/translation="SDAVTKRSASLKWCLPQTHLLGADLLSSITYSERIKIETPCEL
GLSNQMGVLRNRPSTIYSLSRCQPYQAEIAQTDABYKRSQCGRKTKLSDELKQKI
LNLHLSLSPNFRNFKLATKSIYNLQNGRIGFSLNDLPEHGVQRNVRNQSRKNY
QSLGSIIEQRPMFNRNIGIDFELDTVVGPRGHSKAVLLTILDRKSRFLWAYRLKDR
TTATVNEALTKEFLTTPNGPVHSTVDVRGTFSGLVSLBSQYGIKTYCHAYTPADVW
NNALGIYVVFILKGLVLSLVLIKI"
10396..10442
/notes="DR1"
/citation=[1]

```

```

/rpt_type=DIRECT
10397..10443
/notes="IR2"
/citation=[1]
/rpt_type=INVERTED
10681..10716
/standard_name="terminator 2"
/citation=[1]
/complement(10727..11593)
/gene="scrk"
/complement(10727..11593)
/gene="scrk"
/function="fructose 6-phosphotransferase"
/standard_name="fructokinase"
/citation=[1]
/codon_start=1
/transl_table=11
/product="fructokinase"
/protein_id="CAA83667.1"
/db_xref="GI:475967"
/db_xref="GOA:P43468"
/db_xref="Swiss-Prot:P43468"
/translation="MLLGAIBAGTKFVCATGAENGQVSDRISIPITTPVETMTAVDD
YFTHEPDAIGIGSFGPIGVNPHDPKYGIYITTPKPGWGDFFLGHLSQFNIPLYWT
YTVNEAAYGESMIGIAKDVNPSIYMTIGTGAGVISOHFIENGRTHTELGHWRLRL
PGDDFKSNCPYHDIICLEGAAAGPAVKRTGKAGKDI PVDDPWPIITDYLQAQCNLT
VAPAPDKIILNGVMNQRLFPPIREKFPAYLNGYEEVPLDDIIVPAGLNNSGIAG
GLLLQAALKNA"
/complement(13852..15989)
/gene="scrA"
/complement(13852..15989)
/gene="scrA"
/function="enzyme II of the PTS system, sucrose specific"
/standard_name="enzyme IIscr"
/citation=[1]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="enzyme IIabc"
/protein_id="CAA83668.1"
/db_xref="GI:475968"
/db_xref="GOA:P43470"
/db_xref="Swiss-Prot:P43470"
Query Match 18.8%; Score 287.6; DB 1; Length 21839;
Best Local Similarity 58.8%; Pred. No. 4.3e-66;
Matches 539; Conservative 0; Mismatches 364; Indels 13; Gaps 2;
QY 4 ATGGCATCTGCGCGTTGCGGTTCTTGCCAGTGTGTTGGTTTTCACCGCAACCAAGCGT 63
Db 15324 ATGGCTAGTGGCGGTTTACTTCTTGCCAACTCTTTAGGATTTTCAGCAAGAGCGC 15265
QY 64 TTGCGCGCAATAGTTCCTTGGCGCGCGGTATGTGTATGGCGATGGTGTTCGCCAGCTT 123
Db 15264 TTGCGTGGCAACCCCTATCTGGGTGCTACG-ATGGCATGATCATGTTCTTACCACCT 15206
QY 124 GGTGAACGCTAGCAGCTGCGCGCCACCATGGCTGCGGCGGAAATGCCAATGTGTCCT 183
Db 15205 AGTTAACGGCTATAGCGTTTGCACACCATGGCAGCGCGCAAGATGGTTTACTTGAACGT 15146
QY 184 GTTTGGTTTAGATGTTGCCAAGCGGTTTACAGAGCGACCGTGTCTTCCTGCTGCTGGT 243
Db 15145 CTTTGGTTTACGTTTGCAAGCGCGCTATCAAGCCAAAGTGTCTCCAGTCTTAGGCGT 15086
QY 244 TTCTTGGATTTGGCAACGATCGAGAAGTTCCTGCACAGCGACTCAAGGGCACTGCAGA 303
Db 15085 CGCCTTCATTCTAGCTACCTTGAAAAATCTTCCATAAACACATTAAGGGGCAATTCA 15026
QY 304 CTTCTGATCATCTCAGTGTCTGAGTGTCTGCTGCTACCGGATTCCTTCAATTATGCCCAT 363
Db 15025 CTTACGTTTACCCCGATGTTTGCATTTGTGATTAATGTTTCTTAACTTTTACAATCGT 14966
QY 364 TGGCCAGCAATGCGCTGGTGGCGGATGTGCTGGCACACGCTCTACAGGACTTTATCA 423

```







```

Db      4490 TAAATCGGAGAACTTTTCTCACCAGCCGACGAGAAATTAATTATTCCTTATGAAACAGG 4549
QY      1186 CCAATGCTTTTCGCGAGTTCGACCAAGGCTGAGGATGGTTCCAAATGTCGATATCTTGATGCA 1245
Db      4550 TCACGCTTATGTTATTAACCAAAAATGAGGT-----GAGTCTTTTACA 4597
QY      1246 CATTGCTTTCGACACAGTAACTCAACGCGACGCTTTTAACCGCTGAGAAGCAGGG 1305
Db      4598 TATTGGAATAGATACCTGTTTCAATGAATGTAATGATTCATACAAAATGTTAAAGTTGG 4657
QY      1306 CGATGAAGTCAACAGCAGGAGGAGTGTGTGTAATTCGATATTCGATCCCATTAAGGCTGC 1365
Db      4658 CCAGAAAGTAAAGCAGGAGGATTTACTAGATCTTTTGAAGAAGAAATCAAGAAGAG 4717
QY      1366 AGCTTATGAGTAAACACGCGCGATGTTGTTTCGAAT 1403
Db      4718 TGAATGGAATGACTGTAAATTAATTTATTAACAAT 4755

RESULT 15
LOCUS   AE004395/c
DEFINITION Vibrio cholerae O1 biovar eltor str. N16961 linear BCT 10-OCT-2003
ACCESSION AE004395
VERSION   AE003853
KEYWORDS
SOURCE   AE004395.1 GI:9658068
ORGANISM Vibrio cholerae O1 biovar eltor str. N16961
          Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
          Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 10264)
          Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
          Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Unayam, L.,
          Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
          Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
          Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D.,
          Nierman, W.C. and White, O.
          DNA sequence of both chromosomes of the cholera pathogen Vibrio
          cholerae
          Nature 406 (6795), 477-483 (2000)
JOURNAL   20406833
MEDLINE   10952301
PUBMED
REFERENCE 2 (bases 1 to 10264)
          Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
          Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Unayam, L.A.,
          Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
          Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
          Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D.,
          Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
          Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
          Direct Submission
          Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
          Medical Center Dr., Rockville, MD 20850, USA
FEATURES
          Location/Qualifiers
          1..10264
             /organism="Vibrio cholerae O1 biovar eltor str. N16961"
             /mol_type="genomic DNA"
             /strain="N16961"
             /serotype="O1"
             /db_xref="taxon:243277"
             /chromosome="II"
             /note="biotype: El Tor"
             complement (114..1553)
             /gene="VCA0653"
             complement (114..1553)
             /gene="VCA0653"
             /note="similar to GB:M76768 SP:P22825 PID:155262;
             identified by sequence similarity; putative"
             /codon_start=1
             /transl_table=11
             /protein_id="AAF96554.1"

          gene
          CDS

```

```

/db_xref="GI:9658069"
/translation="MDYPVIAKQLLESLOGKSNIQALAHCACTRLRLVLDNTOINESA
ISLSQVKQFKVAGYQIIFGSGIVNQVIAEWAKLTGLVEMSTNDVASAGAKQNSWA
QRAVGLSDIFVRIIPAI VAGGLMGLFNVLTPAGLFIQGLSLIDNPGLADLASMIN
TTANAFVPIVLLAFSAFSGRTGPNFLGAALGMLMVHLDLLNGWFGSGASVSGTPT
WNILAFPEIKVGVQSGVLPLVNSYILAKIENGLRIKIVPSVDNLLTLMIAITGFL
TTTVGLPDRDVGEMLDALNMLYDSAGFVGALFGFIYAPFVITGMHHSFIALETOL
LADIVTGTGTFIFPIAAMSNIAAGRAALAVGWTKETKLKGVAPISGVTALLGTEPA
MGVNLKLRYPPIAICGAALASAFITLNFVKAQALGAAGLPGIISINPOQIYYING
MAISFVAFAFALIVLLAMREKTKQAQVTA"
/gene="VCA0654"
1854..2846
1854..2846
/gene="VCA0654"
/notes="similar to GB:114685 PID:289864; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="sucrose operon repressor Scrr, putative"
/protein_id="AAF96555.1"
/db_xref="GI:9658070"
/translation="WASLHDVARLAGVSKTSVSRVINDEYGVKESKIKVLKAVECGG
YMNQVADKLSQKTNLIGVIVRVSSHATAQGVDLGTAI FQAGKHVILLANTHQVHA
KEIETQIFNQRVEGIVFYATHLQPLVHAIQHSAPVVLVQDQSLNIPISVHDD
HRVFEAGQRLVKAGCKQIFIGVQSDDIADVLRSGQALQALGHQOQTLIPHARGNF
SIESGVMAKQAVLDYFOLDGLFCATDLRAVGAIKALQELGVHVGQVQLGVGNDEL
AVVSNPPISTENYAFDKAGENAAMLLERIAGRQGMKSVLTFQNIQRETCPV"
2322..4568
/gene="VCA0655"
2322..4568
/gene="VCA0655"
/notes="similar to GB:M74035 PID:155269; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="sucrose-6-phosphate dehydrogenase"
/protein_id="AAF96556.1"
/db_xref="GI:9658071"
/translation="MLLTLLLAGINNVTILAPQGVVLAKHPPLVPHLPDDVS
LQSVLGWQLSVQRTAEDSDQOLAIGAIAEROKLETLPYTALDCPYRPLWHISPP
QGLLNDPNGFIYHGEYHLFYQMFACEDKDYVHKLSDLDVQWQSVALTIPSDW
FDSHGVSCHAVSHQOQDLMLFYTNTRGLGVDRQRTQCAARMWANGPEKLGPIVRC
LPEGVTEHIRDPKIVITQCKWMLGAOTLAHOGRLVAVVHSDLLHWHFDELYGDELG
DYGWMECPDMFELQGEAFFVPGQGLASANPHHTIHNRIFRATQNRQAGLALQG
WPLDGFPIYAPQATQADGRVLCGWMGLPDETQHPSCDQWGHQTLALRELEWREG
RIYQPLRELDLQSEPHLTLSDNVTELKTSFALQVTLPGWCELRMLMONTQYRVL
TLDAENQLLRDSRATQIRQGTITRELKDSPTVELRIADQSSLEIFINOQEHVMTS
RIPTPDASGISLHGASVDKLYMAPASAPNLEVNVP"
complement (4690..5661)
/gene="VCA0656"
complement (4690..5661)
/gene="VCA0656"
/notes="similar to GB:M76768 SP:P22824 PID:155263;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="fructokinase"
/protein_id="AAF96557.1"
/db_xref="GI:9658072"
/translation="MNALVRLSSNHLRSDSMSRVLTGDAVVDLIPDQOQHLYKCPG
GAPANVAIARLCGRSAFGRGVNDPFRMQOQTLTDEQVDVCOHLYPDPVHRTSTVV
VDLDEHGRSFTFMVKPSADQFLQSLDIPSQNGEWLHVCSIALANQSRSTFAAIA
QMKEVGVSFEDPNLRVWSEFPQLQATVMRAVGLADVRFSEBELQFLTGTSIEE
GLQIADPQIPLVVVTLGAGLVVTPNSRQVAVKVPDITTTGAGAFVGGLLYRL
SVAQDHNQATILDVAKWANGCAGALATTQKAMTALPNAQALYAFLE"
6008..7546
/gene="VCA0657"
6008..7546
/gene="VCA0657"
/notes="similar to SP:P13035 GB:M21277 GB:M55989 GB:M96795
PID:146181; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11

```

```
/product="aerobic glycerol-3-phosphate dehydrogenase"
/protein_id="AAF96558.1"
/db_xref="GI:9658073"
/translation="MNARSRTQDEVLIIIVGGGNGAGIAADATGRGLKVLGYDA
KFSATSSASRLVHGLLYEYFRLVSEALAESEVLLKKAPHIVTPMRFLLPHR
PFLPAMWIRAGLFLYDNLGRTPSLPASHKVNKAGSVTKPEMOIGFEYSDCWDDAR
LVILNMAQOAGAEVLYNCTVEKAERMGDLHVTLDEOTQOQFERRSHALVNAAG
WVQFLNENAHVSYPYGIIRLIQSGSHIIVPRLHDEPOAYILONEDKRIYFVPIYLLDDYS
MIGTIDVEKGDPRKVAITDARLYLISVNVKHMRELIARDIIAERSGVRLPCLDDSS
NSPOAITRDYTLSDQDAEAPLLSIFGGKUTTYRKLGEAAMKHLAPFFPKMKAPWIA
DAPLGGENFDYALQKLVAAFPFITEVIERWLRSYGSRTQLLAGVGTIEDLGLA
FSGLYQREIDYLCERKFARHAQDIFWRRSLKGLNHDTSVVVEVESYIQQFHAEOPL
KATV"
gene
complement (7648..9258)
/gene="VCA0658"
CDS
complement (7648..9258)
/gene="VCA0658"
/note="similar to GB:X61367 PID:48195; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein"
/protein_id="AAF96559.1"
/db_xref="GI:9658074"
/translation="MYTALIKITLQOAEVSAYTPSAQOEVLVGDHDLVSTDLKG
VITYNDTFPCFIAGQDELKKNENIVRHASMPKAAAFADMMHHLKQGHAWRGIVKNR
TKSGFYWDVAVFYIYQQLTGQSVRVKAEKWEIATYKAYQALLAAKAGKKIQ
FKLHSRLYALLGPMSPALHAPQAEQWNLASRLTLPAGVLGLLFFQELVRIPOQL
KQWNEYDSISRLYSGADAFVADYHLKMASARITILGRMDSAREPLBLANQLHL
TQVEHQALAAQNIQAVTQADAVESAERVSHTSAHQILDOVDHCAETKHSI
NVTHONLQRLATSEASALITLKLSDQAOVQGLMTEIGGIAEOTNLILNAALEAAR
AGEGRGVAVADEBSALTSARTQATQIOTSIDTLMSTIEAWRGDITASRDQTEOCA
QDANTLLOQLQVECVMSDMLRVIGHVASAAHQRELTCENVQNHISIASVATQNSAA
THTVEQLAMNSGKVAERFALSKQPAQK"
gene
complement (9444..10067)
/gene="VCA0659"
CDS
complement (9444..10067)
/gene="VCA0659"
/note="similar to PID:2062640; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="protein F-related protein"
/protein_id="AAF96560.1"
/db_xref="GI:9658075"
/translation="MFEMKFWLVCVTFILLTCGSLVTDPMFNLLDVAPKDPEDVRY
PEWGYAPKNTSVMPNMTQPKRSYEDRLQYLNSGLOHEVVPDGYEMIKLNNTVRF
ETGSARVSAKQWLDTVARFLATSGSDIVLEGHTDNTGSEKLNKLERRANAVKA
ALVQSRVQNAIYTRGFGENVPACTNSTKNGRACNRRVEIRFILASN"
gene
complement (10068..10190)
/gene="VCA0660"
CDS
complement (10068..10190)
/gene="VCA0660"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF96561.1"
/db_xref="GI:9658076"
/translation="MNGEHSNHLCDGKQPICKLLMLTDKVRNMPINYYKFN"
ORIGIN
```

```
Query Match 15.2%; Score 231.8; DB 1; Length 10264;
Best Local Similarity 55.4%; Pred. No. 4.9e-51;
Matches 497; Conservative 0; Mismatches 387; Indels 13; Gaps 2;

QY 3 CATGGCATCTGCGCGTTCGCGTCTTCCAGTGTGGTGGTTTCACCGCAACCAAGCG 62
Db 1071 CTTCCGATGCCCTTTTGTATTGCTGTATTGTTGGCATTTTCTGCTCGCGCA 1012
QY 63 TTTCCGCGCAATGAGTTCCTGGCGCGCGATATGGTATGGGATGGTGTCCCGAGCT 122
Db 1011 GTTTGGCGTAACCCCTTTTGGCGCTGCG-CTGGGTATGCTGATGGTGCATCCCGATC 953
```

Search completed: October 30, 2004, 15:50:51  
Job time : 6350 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 12:39:21 ; Search time 726 Seconds  
(without alignments)  
11041.138 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 ctcattggcattcgcgcgtt.....gttgaaaccttgagtgttcg 1527

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	1527	5	Aaf31528 C.glutami
2	1527	100.0	1527	6	Abx07668 S. pneumo
3	1515	99.2	349980	5	Abx07668 S. pneumo
4	1492	97.7	1983	5	Aah67869 C. glutami
5	1412.6	92.5	5969	4	Aaf32543 Brevibact
6	1107.4	72.5	1109	5	Aaf31529 C.glutami
7	218.8	14.3	110000	10	Continuation (6 of
8	190.6	12.5	2913	4	Aah54100 S. epider
9	177.8	11.6	975	6	Abk73538 Bacillus
10	174.6	11.4	861	8	Acf74550 Staphyloc
11	171.6	11.2	1491	10	Adh84532 Enterococ
12	151.2	9.9	5840	2	Aax12968 Enterococ
13	151.2	9.9	5840	6	Abx07668 S. pneumo
14	143.6	9.4	1386	11	ABD00428
15	143.6	9.4	1452	11	ABD00072
16	137.4	9.0	732	6	ABN91814
17	137.2	9.0	1380	6	Abk73394 Bacillus
18	135.4	8.9	3037	6	Abq70786 Listeria
19	134.6	8.8	110000	6	Continuation (3 of
20	134.6	8.8	110000	6	Abq69245 Listeria
21	132.2	8.7	110000	6	ABA03041 Listeria

22	126.6	8.3	1965	10	ABX07668
23	126.6	8.3	2127	8	ABZ42306
24	126.6	8.3	3895	2	AAV52334
25	126.6	8.3	110000	10	ABS56454_16
26	121.4	8.0	2028	6	ABN66285
27	121.4	8.0	110000	6	ABN71527_01
28	114.6	7.5	951	10	ADH84192
29	113.4	7.4	1884	4	AA555732
30	111	7.3	1543	10	ADD13201
31	109.6	7.2	3615	8	ABZ77360
32	109.6	7.2	3615	10	AAU51848
33	109.4	7.2	1905	10	ADF02562
34	108	7.1	3615	2	AAQ55752
35	107.8	7.1	1185	8	ABZ42298
36	107.8	7.1	9769	2	AAV52163
37	107.8	7.1	110000	10	ABS56454_15
38	105.6	6.9	1014	6	ABK73531
39	103.6	6.8	1917	6	ABN67846
40	103.6	6.8	110000	6	ABN71527_16
41	99	6.5	1098	6	ABK73443
42	98	6.4	489	6	ABK73566
43	97.8	6.4	465	2	AAI47502
44	97.2	6.4	1971	11	ACH95246
45	93.6	6.1	1854	6	ABQ69909

#### ALIGNMENTS

##### RESULT 1

AAF31528  
ID AAF31528 standard; DNA; 1527 BP.  
AC AAF31528;  
XX  
DT 09-APR-2001 (first entry)  
XX  
DE C.glutamicum phosphoenolpyruvate DNA #1.  
XX  
KW Phosphoenolpyruvate; sugar phosphotransferase system; PTS; ds.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN WO200102583-A2.  
XX  
PD 11-JAN-2001.  
XX  
PF 27-JUN-2000; 2000WO-IB000973.  
XX  
PR 01-JUL-1999; 99US-0142691P.  
PR 23-AUG-1999; 99US-0150310P.  
PR 03-SEP-1999; 99DE-01042095.  
PR 03-SEP-1999; 99DE-01042097.  
(BADI ) BASF AG.  
Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
WPI; 2001-080989/09.  
Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:  
sugar phosphotransferase system proteins or their portions, useful for  
typing or identifying C. glutamicum or related bacteria, and as markers  
for transformation.

Claim 3; Page 98-101; 144pp; English.

The present invention relates to Corynebacterium glutamicum phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The PTS nucleic acids and proteins are useful in the identification of microorganisms which can be used to produce fine chemicals, for modulating fine chemical production in C. glutamicum or related bacteria, the typing or identification of C. glutamicum or related bacteria, as

CC reference points for mapping C. glutamicum genome, and as markers for  
CC transformation  
XX  
SQ Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1527; DB 5; Length 1527;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCATGGCAATCGCCCGCTTCCGCTTCTGCGAGTGTTCAGAGTGTGGTTTCCACGCAACCAAG 60  
DB 1 CTCATGGCAATCGCCCGCTTCCGCTTCTGCGAGTGTTCAGAGTGTGGTTTCCACGCAACCAAG 60  
QY 61 CGTTTCGGCGCAATGAGTTCCTGGCGCCGCGATTGCTATGCGATGCGATGCTTCCCGAG 120  
DB 61 CGTTTCGGCGCAATGAGTTCCTGGCGCCGCGATTGCTATGCGATGCGATGCTTCCCGAG 120  
QY 121 CTGTGTGAACGGCTACGAGCTGGCGCCACCATGCTCGCGGCGAAATGCCAATGTGTC 180  
DB 121 CTGTGTGAACGGCTACGAGCTGGCGCCACCATGCTCGCGGCGAAATGCCAATGTGTC 180  
QY 181 CCGTTTGGTTAGATGTTGCCAAGCCGCTTACAGGCGACCGTCTTCCCTGCTGCT 240  
DB 181 CCGTTTGGTTAGATGTTGCCAAGCCGCTTACAGGCGACCGTCTTCCCTGCTGCTGCT 240  
QY 241 GGTTCCTTGGATTCTGGCAACGATCGAGAAGTTCTGTCACAGCGACTCAAGGGCACTGC 300  
DB 241 GGTTCCTTGGATTCTGGCAACGATCGAGAAGTTCTGTCACAGCGACTCAAGGGCACTGC 300  
QY 301 AGACTTCCTGATCACTCCAGTGTCTGACGTTGTCTGCTCACCGGATTCCTTACATTCATCGC 360  
DB 301 AGACTTCCTGATCACTCCAGTGTCTGACGTTGTCTGCTCACCGGATTCCTTACATTCATCGC 360  
QY 361 CATTCGCCAGCAATCGCTGGTGGCGATGCTGTGGCACACGCTCTACAGGCACTTTA 420  
DB 361 CATTCGCCAGCAATCGCTGGTGGCGATGCTGTGGCACACGCTCTACAGGCACTTTA 420  
QY 421 TCATTTCCGTGTCCAGTCCGCGCTGCTCTTCCGTCTGCTCTACTACCAATCGTCAT 480  
DB 421 TCATTTCCGTGTCCAGTCCGCGCTGCTCTTCCGTCTGCTCTACTACCAATCGTCAT 480  
QY 481 CACTGCTCTGCAACCAAGTCTCCCGCAATGAGCTGGAGCTGTGTTAACAGGGTGGATC 540  
DB 481 CACTGCTCTGCAACCAAGTCTCCCGCAATGAGCTGGAGCTGTGTTAACAGGGTGGATC 540  
QY 541 CTTTCATCTCGCAACGGCAATCTATGCTTAATATCGCCCAAGGCTGGGCAATGTTGGCAGT 600  
DB 541 CTTTCATCTCGCAACGGCAATCTATGCTTAATATCGCCCAAGGCTGGGCAATGTTGGCAGT 600  
QY 601 GTTCTCTCGGCAAGAGTGAAGAGCTCAAGGCGCTTGCAAGTGTCTCAGGCTGTCTCCGC 660  
DB 601 GTTCTCTCGGCAAGAGTGAAGAGCTCAAGGCGCTTGCAAGTGTCTCAGGCTGTCTCCGC 660  
QY 661 TGTCTCTGTTATTAACGAGCCTCGCATCTCGGTGTGAACCTTCGCTCGGCTGGCGCTT 720  
DB 661 TGTCTCTGTTATTAACGAGCCTCGCATCTCGGTGTGAACCTTCGCTCGGCTGGCGCTT 720  
QY 721 CTTTCATCGTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATATCAA 780  
DB 721 CTTTCATCGTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATATCAA 780  
QY 781 GGCAGTGTGCTGGCGCTGCAAGTTCCTTGGGTGTGTTCTTATTTGATGCTCCAGATAT 840  
DB 781 GGCAGTGTGCTGGCGCTGCAAGTTCCTTGGGTGTGTTCTTATTTGATGCTCCAGATAT 840  
QY 841 GGTTCATGTTCTGGTGTGCAATGTTTACCTTCTTCATCGCATTCGGCGCAGCGATTGC 900  
DB 841 GGTTCATGTTCTGGTGTGCAATGTTTACCTTCTTCATCGCATTCGGCGCAGCGATTGC 900  
QY 901 TTATGGCCTTTACTTGTTCGCGCAACGCGACATTTGATCCAGATGCAACCGCTGCTCC 960  
DB 901 TTATGGCCTTTACTTGTTCGCGCAACGCGACATTTGATCCAGATGCAACCGCTGCTCC 960

QY 961 AGTCCTCTGCAAGACGACCAAAAGCGAAGCAGACCCCGAGATTTTCAAAGCATTC 1020  
DB 961 AGTCCTCTGCAAGACGACCAAAAGCGAAGCAGACCCCGAGATTTTCAAAGCATTC 1020  
QY 1021 CACCATCATCAGGACCTTTGACGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080  
DB 1021 CACCATCATCAGGACCTTTGACGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080  
QY 1081 CATGTTTCCAGCGAAAGCTTGGCTCGGGCGTTGCCATCGTCCCAACCAAGGGCGAGTT 1140  
DB 1081 CATGTTTCCAGCGAAAGCTTGGCTCGGGCGTTGCCATCGTCCCAACCAAGGGCGAGTT 1140  
QY 1141 AGTTTCTCCGTTGAGTGAAGAATGTTGCTGGCAATTCCTCATCTGSCCATGCTTTCGCACT 1200  
DB 1141 AGTTTCTCCGTTGAGTGAAGAATGTTGCTGGCAATTCCTCATCTGSCCATGCTTTCGCACT 1200  
QY 1201 TCGCACCAAGCTGAGGATGCTTCAATGTGGATATCTTGATGCACTTGGTTTCGACAC 1260  
DB 1201 TCGCACCAAGCTGAGGATGCTTCAATGTGGATATCTTGATGCACTTGGTTTCGACAC 1260  
QY 1261 AGTAAACCTCAACGGCAGCGACTTTAAACCCGCTGAAGAGCAGGCGGATGAAGTCAAAAGC 1320  
DB 1261 AGTAAACCTCAACGGCAGCGACTTTAAACCCGCTGAAGAGCAGGCGGATGAAGTCAAAAGC 1320  
QY 1321 AGGGAGCTGCTGTGTAATTGCAATTCGATGTCCTTAAGCTGAGGTCAGGTTATGAGTAAAC 1380  
DB 1321 AGGGAGCTGCTGTGTAATTGCAATTCGATGTCCTTAAGCTGAGGTCAGGTTATGAGTAAAC 1380  
QY 1381 CACGCGGATGTTGTTTCGAATTACAAGAAACCGGACTGTAAACACTTACGGTTTGGG 1440  
DB 1381 CACGCGGATGTTGTTTCGAATTACAAGAAACCGGACTGTAAACACTTACGGTTTGGG 1440  
QY 1441 CGAAATTTGAAGCGGAGCGCAACCTCTCTCAACGCTCGCAAGAAAGAGCGGTGCCAGCAAC 1500  
DB 1441 CGAAATTTGAAGCGGAGCGCAACCTCTCTCAACGCTCGCAAGAAAGAGCGGTGCCAGCAAC 1500  
QY 1501 ACCATAAGTTGAAACCTTGAAGTGTTCG 1527  
DB 1501 ACCATAAGTTGAAACCTTGAAGTGTTCG 1527  
RESULT 2  
ABS65346  
ID ABS65346 standard; DNA; 1527 BP.  
XX  
AC ABS65346;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE DNA encoding C. glutamicum metabolic pathway (MP) protein #5.  
XX  
KW Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;  
KW cofactor; nucleotide; nucleoside; trehalose; fine chemical production;  
KW organic acid; non-proteinogenic amino acid; purine base; carbohydrate;  
KW pyrimidine base; lipid; unsaturated fatty acid; diols; polyketide;  
KW aromatic compound; food industry; animal feed; cosmetic industry;  
KW pharmaceutical industry; gene; ds.  
XX  
OS Corynebacterium glutamicum ATCC 13032.  
XX  
PN WO200251231-A1.  
XX  
PD 04-JUL-2002.  
XX  
PF 22-DEC-2000; 2000WO-EP013143.  
XX  
PR 22-DEC-2000; 2000WO-EP013143.  
XX  
PA (BADI ) BASF AG.  
XX  
PI Pompejus M, Kroeger B, Zelder O, Schroeder H;  
XX  
XX WPI; 2002-643289/69.

DR P-PSDB; ABG80325.  
XX New metabolic pathway genes of *Corynebacterium glutamicum* for producing  
PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,  
PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical  
PT industries.  
XX  
PS Claim 1; Page 108-111; 176pp; English.  
XX  
CC The present invention relates to the isolation of *Corynebacterium*  
CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide  
CC sequences encoding them. The MP proteins are enzymes involved in the  
CC metabolism of molecules important for the normal functioning of cells  
CC (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or  
CC trehalose). The polynucleotide sequences encoding the MP proteins are  
CC useful for producing fine chemicals, particularly organic acids, non-  
CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,  
CC nucleotides, lipids, (un)saturated fatty acids, diols, carbonhydrates,  
CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The  
CC fine chemicals are useful in the food, animal feed, cosmetic or  
CC pharmaceutical industries. ABS65342-ABS65364 encode the C. glutamicum MP  
CC proteins of the invention  
XX  
SQ Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1527; DB 6; Length 1527;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCATGGCATCTGCGCGTTCGGTCTTTCGACAGTGTGGTTGGTTTACCGCAACAAG 60  
DB 1 CTCATGGCATCTGCGCGTTCGGTCTTTCGACAGTGTGGTTGGTTTACCGCAACAAG 60  
QY 61 CGTTTCGGCGGCAATCAGATTCTCGGGCGCGGTATTGGTATGCGATGCTGTTCCCGAG 120  
DB 61 CGTTTCGGCGGCAATCAGATTCTCGGGCGCGGTATTGGTATGCGATGCTGTTCCCGAG 120  
QY 121 CTTGGTGAACGGCTACGACGTGCGCGCCACCATGGCTGGCGGCAATGCCAATGTGGTC 180  
DB 121 CTTGGTGAACGGCTACGACGTGCGCGCCACCATGGCTGGCGGCAATGCCAATGTGGTC 180  
QY 181 CTTGTTGGTTTATGTTGGCCCAACCGGTTACCAAGGCAACCGTCTTCCTGTGTGGT 240  
DB 181 CTTGTTGGTTTATGTTGGCCCAACCGGTTACCAAGGCAACCGTCTTCCTGTGTGGT 240  
QY 241 GGTTCCTTGATTTCTGCAACGATCAGAGTTCTCTGCAACAGCACTCAAGGGCACTGC 300  
DB 241 GGTTCCTTGATTTCTGCAACGATCAGAGTTCTCTGCAACAGCACTCAAGGGCACTGC 300  
QY 301 AGACTTCCTGATCACTCCAGTGTGACGTTGCTGCTCACCGGATTCCTTACATTCATCGC 360  
DB 301 AGACTTCCTGATCACTCCAGTGTGACGTTGCTGCTCACCGGATTCCTTACATTCATCGC 360  
QY 361 CATTGGCCCAAGATCGCTGGTGGCGCATGTCTGGCACACAGCTCTACAGGCACTTTA 420  
DB 361 CATTGGCCCAAGATCGCTGGTGGCGCATGTCTGGCACACAGCTCTACAGGCACTTTA 420  
QY 421 TGAATTCGGTGTCTCAGTCCGGGTCTGTCTCTTCGGTCTGTCTACTCACCAATCGTCAT 480  
DB 421 TGAATTCGGTGTCTCAGTCCGGGTCTGTCTCTTCGGTCTGTCTACTCACCAATCGTCAT 480  
QY 481 CACTGGTCTGCAACCAATCTCCCGCAATGAGCTGGAGCTGTTTAAACAGGGTGGATC 540  
DB 481 CACTGGTCTGCAACCAATCTCCCGCAATGAGCTGGAGCTGTTTAAACAGGGTGGATC 540  
QY 541 CTTTCATCTTCGCAACGGCATCTATGCTAATATCGCCCGAGGGTGGCGCATGTTTGGCAGT 600  
DB 541 CTTTCATCTTCGCAACGGCATCTATGCTAATATCGCCCGAGGGTGGCGCATGTTTGGCAGT 600  
QY 601 GTTCTCTTCGCGGCAAGAGTGAAGCTCAAGGCTTTCAGGCTTTCAGGTTCTCCCGC 660  
DB 601 GTTCTCTTCGCGGCAAGAGTGAAGCTCAAGGCTTTCAGGCTTTCAGGTTCTCCCGC 660

QY 661 TGTTCCTTGGTATTACCGAGCTCGCATCTTCGGTGTGAACCTTCGCGCTGGCGGT 720  
DB 661 TGTTCCTTGGTATTACCGAGCTCGCATCTTCGGTGTGAACCTTCGCGCTGGCGGT 720  
QY 721 CTTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCGATCTTTAATCAA 780  
DB 721 CTTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCGATCTTTAATCAA 780  
QY 781 GGCAGTTCGGTTCGGCGCTGCGAGTTCTTGGTGTGTTCTTATGATCTCCAGATAT 840  
DB 781 GGCAGTTCGGTTCGGCGCTGCGAGTTCTTGGTGTGTTCTTATGATCTCCAGATAT 840  
QY 841 GGTTCATGTTCTTCGGTGTGCGAGTTCTTCTTCATCGCATTCGGCGCAGCATTCG 900  
DB 841 GGTTCATGTTCTTCGGTGTGCGAGTTCTTCTTCATCGCATTCGGCGCAGCATTCG 900  
QY 901 TTATGCGCTTTACTTGGTTCGGCGCAACGGCAGCATTCGATCGAGTGCACACCGCTGCTC 960  
DB 901 TTATGCGCTTTACTTGGTTCGGCGCAACGGCAGCATTCGATCGAGTGCACACCGCTGCTC 960  
QY 961 AGTGCCTGCGAGCAACGACCAAGCCGAGCAGNAGCACCCGCGCAGATTTTCAACGATTC 1020  
DB 961 AGTGCCTGCGAGCAACGACCAAGCCGAGCAGNAGCACCCGCGCAGATTTTCAACGATTC 1020  
QY 1021 CACCATCATCCAGGCACTTTTGACCGGTCAAGCTATTGCACTCAGCAGCGTCAGCGATGC 1080  
DB 1021 CACCATCATCCAGGCACTTTTGACCGGTCAAGCTATTGCACTCAGCAGCGTCAGCGATGC 1080  
QY 1081 CATGTTTGCAGCGGAAAGCTTGGCTCGGGCGTTCGATCGTCCCAACCAAGGGCGAGTT 1140  
DB 1081 CATGTTTGCAGCGGAAAGCTTGGCTCGGGCGTTCGATCGTCCCAACCAAGGGCGAGTT 1140  
QY 1141 AGTTTCTCCGCTGAGTGGAAAGATTGTGTGGCATTCCTCATCTGGCGCATGCTTCCAGT 1200  
DB 1141 AGTTTCTCCGCTGAGTGGAAAGATTGTGTGGCATTCCTCATCTGGCGCATGCTTCCAGT 1200  
QY 1201 TCGCACCAAGGCTGAGGATGTTCCAAATGTGGATATCTTTGATGCATTCATTTGTCACAC 1260  
DB 1201 TCGCACCAAGGCTGAGGATGTTCCAAATGTGGATATCTTTGATGCATTCATTTGTCACAC 1260  
QY 1261 AGTAAACCTCAACGGCAGCACCTTTTAAACCGCTGGAAGAAGCGGCGATGAAGTCAAAGC 1320  
DB 1261 AGTAAACCTCAACGGCAGCACCTTTTAAACCGCTGGAAGAAGCGGCGATGAAGTCAAAGC 1320  
QY 1321 AGGGGAGCTGCTGTGTAATTCGATATTCGATTCGATTAAGGCTGAGGTTATGAGTAAAC 1380  
DB 1321 AGGGGAGCTGCTGTGTAATTCGATATTCGATTCGATTAAGGCTGAGGTTATGAGTAAAC 1380  
QY 1381 CACGCGGATTTGTTTTCGAATTAACAAGAAACCGGACCTGTAAACACTTACGGTTTGGG 1440  
DB 1381 CACGCGGATTTGTTTTCGAATTAACAAGAAACCGGACCTGTAAACACTTACGGTTTGGG 1440  
QY 1441 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACGCTCGCAAGAAAGAGCGGTGCCAGCAAC 1500  
DB 1441 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACGCTCGCAAGAAAGAGCGGTGCCAGCAAC 1500  
QY 1501 ACCATAAGTTGAACCTTGAGTGTTCG 1527  
DB 1501 ACCATAAGTTGAACCTTGAGTGTTCG 1527

## RESULT 3

AAH68533/c

ID AAH68533 standard; DNA; 349980 BP.

XX AAH68533;

XX AC

XX DT 26-SEP-2001 (first entry)

XX C

DE C glutamicum coding sequence fragment SEQ ID NO: 7068.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

KW

[illegible]

[illegible]

```
QY 1261 AGTAACTCAACGCGACGACTTTAACCCCTGAAGAAGCAGGCGGATGAAGTCAAAAGC 1320
DB |||||
QY 1740 AGTAACTCAACGCGACGACTTTAACCCCTGAAGAAGCAGGCGGATGAAGTCAAAAGC 1799
DB |||||
QY 1321 AGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGCTGCGAGGTTATGAGGTAAC 1380
DB |||||
QY 1800 AGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGCTGCGAGGTTATGAGGTAAC 1859
DB |||||
QY 1381 CACGCCGATGTTGTTTCGAATTCAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440
DB |||||
QY 1860 CACGCCGATGTTGTTTCGAATTCAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1919
DB |||||
QY 1441 CGAAATTAAGCGGAGCGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCCAGCAAC 1500
DB |||||
QY 1920 CGAAATTAAGCGGAGCGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCCAGCAAC 1979
DB |||||
QY 1501 ACCA 1504
DB |||||
QY 1980 ACCA 1983
DB |||||
RESULT 5
AAF32543
ID AAF32543 standard; DNA; 5969 BP.
XX
AC AAF32543;
XX
DT 11-SEP-2003 (revised)
DT 20-APR-2001 (first entry)
XX
XX Brevibacterium lactofermentum sucrose PTS enzyme II DNA SEQ ID NO:1.
XX
XX Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;
KW phosphoenolpyruvate:carboxylate phosphotransferase system; glucose;
KW coryneform bacterium; phosphoenolpyruvate-sugar transport system; ds.
XX
XX Corynebacterium glutamicum.
XX
XX Key Location/Qualifiers
FT CDS 3779..5764
FT /*tag= a
FT /product= "sucrose PTS enzyme II"
XX
XX WO200102584-A1.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-JP004348.
XX
XX 02-JUL-1999; 99JP-00189512.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;
XX
XX WPI: 2001-138150/14.
XX P-PDB; AAB69080.
XX
XX Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme
XX II obtained by cassette ligation-mediated amplification of downstream
XX domain of coryneform bacterium sucrose gene, with sucrose-binding
XX activity.
XX
XX Claim 3; Page 22-29; 45pp; Japanese.
XX
XX The present sequence encodes the Brevibacterium lactofermentum sucrose
XX PTS (phosphoenolpyruvate:carboxylate phosphotransferase system or
XX phosphoenolpyruvate-sugar transport system) enzyme II, which has sucrose-
XX binding activity. A coryneform bacterium produced with the sucrose PTS
XX enzyme II gene can have more efficient sugar uptake, and improved amino-
XX acid and nucleic acid productivity. The sucrose PTS gene and it's
XX disrupted gene, such as one without the sucrose PTS function, can be used
XX to produce new breeds of coryneform bacterial strains to uptake sugar
```

```
CC more efficiently e.g. glucose only or and sucrose, and can have improved
CC amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 5969 BP; 1388 A; 1575 C; 1578 G; 1428 T; 0 U; 0 Other;
Query Match 92.5%; Score 1412.6; DB 4; Length 5969;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;
QY 1 CTCATGGCATCTGGCCGCTTCGGCTTCCTGCGGCGCGGATTCGATGGGATGGTGTTCGCCAG 60
DB CTGATGGCATCTGGCCGCTTCGGCTTCCTGCGGCGCGGATTCGATGGGATGGTGTTCGCCAG 4318
QY 61 CGTTTCGCGCAATGAGTTCCTGCGGCGCGGATTCGATGGGATGGTGTTCGCCAG 120
DB CGTTTCGCGCAATGAGTTCCTGCGGCGCGG- GCAATTGATGGGATGGTGTTCGCCAG 4377
QY 121 CTTCGTGAACGGCTACGACGTCGGCGCCACCATGCTCGGGCGGAAATGCCAATGTGTC 180
DB CTTCGTGAACGGCTACGACGTCGGCGCCACCATGACCGCGGCGGAAATGCCAATGTGTC 4437
QY 181 CCTGTTTGGTTAGATGTTGCCAAGCCGGTTACAGGGCACCGTCTTCCTGTCGTGT 240
DB CCTGTTTGGTTAGATGTTGCTCAAGCTGGTTACAGGGCGCACCGTCTTCCTGTCGTGT 4497
QY 241 GGTTCCTTGGATTCGCGCAACGATCGAAGATTCCTGCAAGGACACTCAAGGGCACTGC 300
DB GGTTCCTTGGATTCGCGCAACGATCGAAGATTCCTGCAAGGACACTCAAGGGCACTGC 4557
QY 301 AGACTTCCTGATCACTCCAGTCTGACGTTGCTGTCTCACCGGATTCCTTACATTCATCGC 360
DB AGACTTCCTGATCACTCCAGTCTGACGTTGCTGTCTCACCGGATTCCTTACATTCATTCG 4617
QY 361 CATTCGCCAGCAATGCGCTGGTGGGATGTCGTGACACCGCTCTACAGGGACTTTA 420
DB CATTCGCCAGCAATGCGCTGGTGGGATGTCGTGACACCGCTCTACAGGGACTTTA 4677
QY 421 TGATTTTCGTTGTCAGTCGGCGGTCGTCTCTTCGTTCTGTTCTACTCACCAATGCTCAT 480
DB TGATTTTCGTTGTCAGTCGGCGGTCGTCTCTTCGTTCTGTTCTACTCACCAATGCTCAT 4737
QY 481 CACTGCTCTGCACCACTCCTCCGCCAATTCAGCTGGAGCTGTTTAAACAGGGTGATC 540
DB CACTGCTCTGCACCACTCCTCCGCCAATTCAGCTGGAGCTGTTTAAACAGGGTGATC 4797
QY 541 CTTTCATCTTCGCAACGGCATCTATGGCTAATATCCGCCAGGCTCGGCGATGTTTGGCACT 600
DB CTTTCATCTTCGCAACGGCATCTATGGCTAATATCCGCCAGGCTCGGCGATGTTTGGCACT 4857
QY 601 GTTCTTCCTGGCGAAGAGTGAAGCTCAAGGGCTTCGAGGTGCTTCAGGTGCTCTCCGC 660
DB GTTCTTCCTGGCGAAGAGTGAAGCTCAAGGGCTTCGAGGTGCTTCAGGTGCTCTCCGC 4917
QY 661 TGTTCCTGTTATTCAGGACCTTCGATCTTCGGTGTGAACCTTCGCTCGCTCGGCTGGCT 720
DB TGTTCCTGTTATTCAGGACCTTCGATCTTCGGTGTGAACCTTCGCTCGCTCGGCTGGCT 4977
QY 721 CTTTCATCGTATTCGTTACCGGACTATCGTGGCGCTTTGATTGCACTCTTTAATCAATCAA 780
DB CTTTCATCGTATTCGTTACCGGACTATCGTGGCGCTTTGATTGCACTCTTTAATCAATCAA 5037
QY 781 GGCAGTTGCGTTGGCGCTGCGAGGTTTCTTGGGTGTGTTTCTATTGATGCTCCAGATAT 840
DB GGCAGTTGCGTTGGCGCTGCGAGGTTTCTTGGGTGTGTTTCTATTGATGCTCCAGATAT 5097
QY 841 GGTTCATGTTTGGTGTGCGATGTTTACCTTCTTCATCGCATTCGGCGGAGCGATTCG 900
DB GGTTCATGTTTGGTGTGCGATGTTTACCTTCTTCATCGCATTCGGCGGAGCGATTCG 5157
QY 901 TTATGCGCTTACTTGGTTTCGGCGCAACGCGAGCATTTGATCCAGATGCAACCGCTGCTCC 960
DB TTATGCGCTTACTTGGTTTCGGCGCAACGCGAGCATTTGATCCAGATGCAACCGCTGCTCC 5217
```

961 AGTGCTGACGAGAACACCAAGCCGAGAGCAGACCCGACGAAATTTTCAAAACGATTC 1020  
5218 AGTGCTGACGAGAACACCAAGCCGAGAGCAGACCCGACGAAATTTTCAAAACGATTC 5277  
1021 CACCATATCCAGGACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080  
5278 CACCATATCCAGGACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 5337  
1081 CATGTTTGCAGCGGAAAGCTTGCCTGGGCGTTGCCATCGTCCCAACCAAGGGGAGTT 1140  
5338 CATGTTTGCAGCGGAAAGCTTGCCTGGGCGTTGCCATCGTCCCAACCAAGGGGAGCT 5397  
1141 AGTTCTCCGGTGAGTGGAAAGATTGTTGGTGGCAATTCCTGCGGCATCTTTTCGCAGT 1200  
5398 GGTTCACCAAGTGAGCGGAAAGATCGTGGTGGCTTCCCAATCTGGTCAAGCTTTCGCAGT 5457  
1201 TGCACCAAGCTGAGATGGTTCCAAATGTGGATATCTTGATGCAATGTTTTCGACAC 1260  
5458 CCGCACTAAGGCTGAGATGGTTCCAAATGTGGATATCTTGATGCAATGTTTTCGACAC 5517  
1261 AGTAAACCTCAACCGCACGCACTTTAAACCGCTGAAGAGCAGGCGATGAAGTCAAAGC 1320  
5518 CGTAAACCTCAACCGCACGCACTTTAAACCGCTGAAGAGCAGGCGATGAAGTCAAAGC 5577  
1321 AGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGGCTGCAGGTTATGAGGTAAC 1380  
5578 AGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGGCTGCAGGTTATGAGGTAAC 5637  
1381 CAGCGCGATTGTTTTCGAAATTAACAGAAACCGGACCTGTAAACACATTTACGGTTTGGG 1440  
5638 CAGCGCGATTGTTTTCGAAATTAACAGAAACCGGACCTGTAAACACATTTACGGTTTGGG 5697  
1441 CGAAATTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCCACCAAC 1500  
5698 CGAAATTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCCACCAAC 5757  
1501 ACATAAGTTGAACCTTGAGTGTTCG 1527  
5758 ACCATAAGTTGAACCTTGAGTGTTCG 5784

## RESULT 6

AAF31529  
ID AAF31529 standard; DNA; 1109 BP.

XX AAF31529;

XX 09-APR-2001 (first entry)

XX C.glutamicum phosphoenolpyruvate DNA #2.

XX Phosphoenolpyruvate; sugar phosphotransferase system; PTS; ds.

XX Corynebacterium glutamicum.

XX WO200102583-A2.

XX 11-JAN-2001.

XX 27-JUN-2000; 2000WO-IB000973.

XX 01-JUL-1999; 99US-0142691P.

XX 23-AUG-1999; 99US-0150310P.

XX 03-SEP-1999; 99DE-01042095.

XX 03-SEP-1999; 99DE-01042097.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habershauer G;

XX WPI; 2001-080989/09.

XX

PT Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:  
PT sugar phosphotransferase system proteins or their portions, useful for  
PT typing or identifying C. glutamicum or related bacteria, and as markers  
PT for transformation.

XX Claim 5; Page 103-104; 144pp; English.

XX The present invention relates to Corynebacterium glutamicum  
CC phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The  
CC PTS nucleic acids and proteins are useful in the identification of  
CC microorganisms which can be used to produce fine chemicals, for  
CC modulating fine chemical production in C. glutamicum or related bacteria,  
CC the typing or identification of C. glutamicum or related bacteria, as  
CC reference points for mapping C. glutamicum genome, and as markers for  
CC transformation

XX Sequence 1109 BP; 236 A; 278 C; 303 G; 292 T; 0 U; 0 Other;

Query Match 72.5%; Score 1107.4; DB 5; Length 1109;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 419 TATGATTTCCGGTCCAGTCGCGGCTCTCTCTTCGCTCTGCTCTACTCACAATCGTC 478

Db 1 TATGATTTCCGGGCTCCAGTCGCGGCTCTCTCTTCGCTCTGCTCTACTCACAATCGTC 60

QY 479 ATCACTGGTCTGCACAGCTCTTCCGCCAATTGAGCTGAGCTGTTTAAACAGGGTGA 538

Db 61 ATCACTGGTCTGCACAGCTCTTCCGCCAATTGAGCTGAGCTGTTTAAACAGGGTGA 120

QY 539 TCCTTCATCTTCGCAACGCACTATGCTTAATATCGCCAGGCTCGGCATCTTTGGCA 598

Db 121 TCCTTCATCTTCGCAACGCACTATGCTTAATATCGCCAGGCTCGGCATCTTTGGCA 180

QY 599 GTGTTCTTCCTGCGAAGAGTGAAAAGCTCAAGGGCTTTCAGGCTCTTCAGGTCTCC 658

Db 181 GTGTTCTTCCTGCGAAGAGTGAAAAGCTCAAGGGCTTTCAGGCTCTTCAGGTCTCC 240

QY 659 GCTGTTCTTGTGTTATACGAGGCTCGCATCTTCGGTGTGAACCTTCGCTCGCTGGCG 718

Db 241 GCTGTTCTTGTGTTATACGAGGCTCGCATCTTCGGTGTGAACCTTCGCTCGCTGGCG 300

QY 719 TTCTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCACCTCTTTAATATC 778

Db 301 TTCTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCACCTCTTTAATATC 360

QY 779 AAGGAGTTGCGTTGGCGCTCCAGGTTTCTTGGGTGTTGTTTCTATTGATGCTCCAGAT 838

Db 361 AAGGAGTTGCGTTGGCGCTCCAGGTTTCTTGGGTGTTGTTTCTATTGATGCTCCAGAT 420

QY 839 ATGCTCATGTTCTTGGTGTGTCAGTTGTTACTTCTTTCATCGCATTCGGCGAGCGATT 898

Db 421 ATGCTCATGTTCTTGGTGTGTCAGTTGTTACTTCTTTCATCGCATTCGGCGAGCGATT 480

QY 899 GCTTATGGCTTTTACTTGGTTCGCGCAACCGCAGCATTGATCCAGATCAACCGCTGCT 958

Db 481 GCTTATGGCTTTTACTTGGTTCGCGCAACCGCAGCATTGATCCAGATCAACCGCTGCT 540

QY 959 CCAGTGCCTGCGAGGAAACGACAAAGCCGAAAGCAGAGCAACCCGCAAGATTTTCAACAGAT 1018

Db 541 CCAGTGCCTGCGAGGAAACGACAAAGCCGAAAGCAGAGCAACCCGCAAGATTTTCAACAGAT 600

QY 1019 TCCACCATCATCCAGGCACTTTGACCGGTGAGCTATTGCACTGAGCAGCGCTCAGCGAT 1078

Db 601 TCCACCATCATCCAGGCACTTTGACCGGTGAGCTATTGCACTGAGCAGCGCTCAGCGAT 660

QY 1079 GCATGTTTTCGAGCGGAAAGCTTCGCTCGCGGCTTGCCTCGTCCCAACCAAGGGGCGAG 1138

Db 661 GCATGTTTTCGAGCGGAAAGCTTCGCTCGCGGCTTGCCTCGTCCCAACCAAGGGGCGAG 720

QY 1139 TTAGTTTCTCCGGTGAGTGGAAAGATTGTGTGGCAATTCCTCATCTGGCCATGCTTTCGCA 1198

Db 721 TTAGTTTCTCCGGTGAGTGGAAAGATTGTGTGGCAATTCCTCATCTGGCCATGCTTTCGCA 780

Qy	1199	GTTCCGACCAAGCTGAGANGTGTTCGAATGTGGATATCTTGATGCACATTTGGTTTCGAC	1258
Dd	781	GTTCCGACCAAGGCTGAGGATCGTTTCCAATGTGGATATCTTGATGCACATTTGGTTTCGAC	840
Qy	1259	ACAGTAAACCTCAACGGCAGCACATTAAACCGCTGAAGAAGCAGGCGCATGAAGTCAA	1318
Dd	841	ACAGTAAACCTCAACGGCAGCACATTAAACCGCTGAAGAAGCAGGCGCATGAAGTCAA	900
Qy	1319	GCAGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGCTGCAGGTTATGAGGTA	1378
Dd	901	GCAGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGCTGCAGGTTATGAGGTA	960
Qy	1379	ACCAGCCGATTTGTTTTCGAATTCACAAGAAAACCGGACCTCTAAACACTTACGGTTTG	1438
Dd	961	ACCAGCCGATTTGTTTTCGAATTCACAAGAAAACCGGACCTCTAAACACTTACGGTTTG	1020
Qy	1439	GCGCAAAATTGAAGCGGGAGCCAACTGCTCAACGTCGCAAGAAAAAGACGGTGCCAGCA	1498
Dd	1021	GCGCAAAATTGAAGCGGGAGCCAACTGCTCAACGTCGCAAGAAAAAGACGGTGCCAGCA	1080
Qy	1499	ACACCATAAGTTGAACCACTTGAGTGTTG	1527
Dd	1081	ACACCATAAGTTGAACCACTTGAGTGTTG	1109

RESULT 7  
ADF77343\_05  
Continuation (6 of 20) of ADF77343 from base 500001 (lactic acid bacteria Lactobacillus reuteri strain ATCC 55969)  
WP Sequence split into 20 fragments LOCUS ADF77343 Accession ADF77343

WP	Fragment Name	Begin	End
WP	ADF77343_00	1	110000
WP	ADF77343_01	100001	210000
WP	ADF77343_02	200001	310000
WP	ADF77343_03	300001	410000
WP	ADF77343_04	400001	510000
WP	ADF77343_05	500001	610000
WP	ADF77343_06	600001	710000
WP	ADF77343_07	700001	810000
WP	ADF77343_08	800001	910000
WP	ADF77343_09	900001	1010000
WP	ADF77343_10	1000001	1110000
WP	ADF77343_11	1100001	1210000
WP	ADF77343_12	1200001	1310000
WP	ADF77343_13	1300001	1410000
WP	ADF77343_14	1400001	1510000
WP	ADF77343_15	1500001	1610000
WP	ADF77343_16	1600001	1710000
WP	ADF77343_17	1700001	1810000
WP	ADF77343_18	1800001	1910000
WP	ADF77343_19	1900001	1983043

RESULT 7

ADP77343\_05  
Continuation (6 of 20) of ADP77343 from base 500001 (lactic acid bacteria *Lactobacillus*  
wp Sequence split into 20 fragments LOCUS ADP77343 Accession ADP77343

WP	sequence split name	Begin	End	Accession
WP	ADFF77343_00	1	110000	ADL77343
WP	ADFF77343_01	100001	210000	
WP	ADFF77343_02	200001	310000	
WP	ADFF77343_03	300001	410000	
WP	ADFF77343_04	400001	510000	
WP	ADFF77343_05	500001	610000	
WP	ADFF77343_06	600001	710000	
WP	ADFF77343_07	700001	810000	
WP	ADFF77343_08	800001	910000	
WP	ADFF77343_09	900001	1010000	
WP	ADFF77343_10	1000001	1110000	
WP	ADFF77343_11	1100001	1210000	
WP	ADFF77343_12	1200001	1310000	
WP	ADFF77343_13	1300001	1410000	
WP	ADFF77343_14	1400001	1510000	
WP	ADFF77343_15	1500001	1610000	
WP	ADFF77343_16	1600001	1710000	
WP	ADFF77343_17	1700001	1810000	
WP	ADFF77343_18	1800001	1910000	
WP	ADFF77343_19	1900001	1983043	

81920	Db	TTATACAAACAACATGGTGCCCTTTGGTATGGGTATCTTTGGTTTATCTTACTCAGCTATT	81979
476	Qy	GTCATCACTGGTCTGCACACAGTCCCTCCCGCCAAATTGAGCTGAGAGCTGTTTAAACAGGGT	535
81980	Db	GTTACTACTGGTCTTCCAAAGATTTCCAGCAGTTGAAACTCAACTTTTAGCTGAATAT	82039
536	Qy	GGATC-----CTTCATCTTCGCAACGCCATCTATGGCTAATATCGCC	577
82040	Db	GCACTGGTCGTGGATCCGGTGCACTTATCTTTGTTCACGCATGTATGCTAAACGTGCT	82099
578	Qy	CAGGTGGGATGTTTGGCAGTGTCTTCCTCGGCAAGAGTGAAGAGCTCAAGAGGCCCTT	637
82100	Db	CAAGTGTCTACATTTGCTATTTACTTCTTAACTAAGAAAGAAAGTTAAAGGTTTA	82159
638	Qy	GCAGTGTCTCAGGTGTCTCCGCTGTTCTTTGGTATTAACGGAGCCTCGGATCTTCGGTGTG	697
82160	Db	GCAAGTTCACTGGTGTGTTCTGCATTACTTGGTATTAACCGAAGCAGCATTTGTTGGTGT	82219
698	Qy	AACCTTTCGCTCGGCTGGCGGTCTTTCATCGGTATCGGTACCGCAGCATTCGCTGGCGCT	757
82220	Db	AACTTGAAGTACAAGTTCCTCATCTTCTGTGCTTTAATTGGTGCAAGTGTGGTGTGCT	82279
758	Qy	TTGATTTGCACCTTTTAAATATCAAGCAGTTGGTTCGGCGCTGCAGGTTTCTTGGGTGTT	817
82280	Db	TTTGCAGGATTAATGCAATGTTACTGCTGCTCTTGGTTCAGCTGGTTCCTTGGTTTC	82339
818	Qy	GTTTCTATTATGATGCTCCAGATATGGTCATGTTCTTGGTGTGTCAGTGTGTACCTTCTTC	877
82340	Db	TTATCAATGGTGCCAAAGAGCATTCAAATGTGGGCATCTCGGTGTGTAATTTAGTTTATT	82399
878	Qy	ATCGCATTCGGCGCAGCGATGCTTATGGCTTTTACTTGGTTCGCGCAACGGCAGCATT	937
82400	Db	GTTTGTCTTTGGTTTAACCTTTGTATACGGTTAAACGTCACCTTTAAGAAGACGTAGTTGAA	82459
938	Qy	GATCCAGATGCAACCGCTGCTCAGTGCTGCAGGAAAGCAACCAAGCCGAAGCAGAGCA	997
82460	Db	GAATCTGTGTACTGTTGAATCTGTGTGTATCAAGTAGCAACAACAAGAAAAGCTGAAACA	82519
998	Qy	CCCGCAGAAATTTCAAACGATTCACCA-----TCATCCAGGCACCTTTGACCGGTGAA	1051
82520	Db	ATTATCAAGAAGATTAAGAATTAATGATGAATTAATGCTCTCTGTTTCAGGTAAA	82579
1052	Qy	GCTATTGCACTGAGCAGCGTCAAGCATGCCATGTTTCCAGCGGAAGCTTGGCTCGGGC	1111
82580	Db	GCAGAAAAGCTTGAAAGATGTAAATGACCCAGTCTTCTACTGAAAGCAATGGTAAAGGT	82639
1112	Qy	GTTGCCATCGTCCCAACCAAGGGCAGTTAGTTTCTCCGCTGAGTGAAGAGATTTGCTG	1171
82640	Db	GCTGCAATGGTACCAAGTGAAGTACTATCTATCCTCCAGTAACTGGTGAATTTACAGTT	82699
1172	Qy	GCATTCCTCATCTGGCCATGCTTTCCAGTTCGACCAAGGCTGAGGATGGTTCCAATGTG	1231
82700	Db	GCTTATGAACATAAGCATGCTTATGGAATTAATTCAGACGAAG-----GTGCA	82747
1232	Qy	GATATCTTGAATGACATTTGGTTTCGACACAGTAAACCTCAACGGCAGCATTTTAAACCG	1291
82748	Db	GAAGTCTTAATTCATATGGTATCGATCTGTAAACATGAAGGGTGAAGGATTTTACTACT	82807
1292	Qy	CTGAAGACGAGGGCGATCAAGTCAAGCAGAGGGGAGCTGCTGTGTGAATTCGATTTGAT	1351
82808	Db	GATGTAAACAAGGCCACAATGTTGAAGGGTGAAGAACTTGGACAGTTGATTTAGAT	82867
1352	Qy	GCCATTAAAGGCTCAGGTTATAGGTAACCAACGCCGATTTGTTTTCGAATTAACAAGAA	1411
82868	Db	GCTGTAAAGAAAGCAGGATACGATCAACTGTCTATGCTGTAATTTACTAACACTACTAGC	82927
1412	Qy	ACCGACCTGTAAACACTTACGTTTGGGCGAAATTTGAAGCGGGAGGCAACCTGCTCAAC	1471
82928	Db	TATGCAAAATGTTCAAGTATTTGATGGCTAGAAAGAGCAATGTTGACGACTTAATCGCA	82987
1472	Qy	GTCCGCAAGAAAGAGCGGTGCCAGCAACACCATAA	1507

```
Db 82988 GTTACTAAGAGATAATTATTGATAAGTAAATCAAAA 83023
RESULT 8
AAH54100/c
ID AAH54100 standard; DNA; 2913 BP.
XX AC AAH54100;
XX DT 03-SEP-2001 (first entry)
XX DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3464.
XX KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
XX OS endocarditis; ds.
XX PN Staphylococcus epidermidis.
XX PD WO200134809-A2.
XX PF 17-MAY-2001.
XX PR 09-NOV-2000; 2000WO-US030782.
XX PR 09-NOV-1999; 99US-0164258P.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Kimmerly WJ;
XX PI WPI; 2001-316495/33.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX PS useful for vaccinating against infections, e.g. endocarditis.
XX PS Claim 8; Page 1030-1031; 2188pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX SQ
XX Sequence 2913 BP; 1072 A; 480 C; 419 G; 942 T; 0 U; 0 Other;
Query Match 12.5%; Score 190.6; DB 4; Length 2913;
Best Local Similarity 55.0%; Pred. No. 1.7e-45;
Matches 468; Conservative 0; Mismatches 364; Indels 19; Gaps 4;
XX 5 TCGCATCTCGCGGTTCGGTCTTCCAGTGTGGTTCACCGCAACCAAGCGTT 64
XX 2690 TTGCAATCCACTTTTACCAATACTTATTGGTTTATAGTGCAGTAAAGCGAT 2631
XX 65 TCGGCGGCAATGAGTTCCTCGGCGCGCGTATTTGGTATGGCGATGTGTTCGAGCTTG 124
XX 2630 TTGTGGTAACTCTTATTAGTGGAGC-TCTAGGTATGATGACTTGTTCATCTGGAATG 2572
XX 125 GTGAACGGGTACGAGGTGCCCGCCACCATGGTGGCGGGCGAA---ATGCCAATGTGGTCC 181
```

```
Db 2571 ATGAGTGCAATGATTTTCCAAAAGCTTTAGAGAAGAAAGCTATTCACACTGGAT 2512
QY 182 CTGTTTGGTTTAGATGTGCCCCAAGCGGTTACAGGGCACCGTGTCTTCTGTGCTGGTG 241
Db 2511 GTCTTTGGGCTACATATTAATGAAGTAGGTTATCAGGGACAAGTATTACCTATGCTCGTA 2452
QY 242 GTTTCTTGGATCTCGCAAGCATCGAGAAGTTCCTGCACAAGCGACTCAAGGGCACTGCA 301
Db 2451 GCAACATATATTTTAGCTAGATTGAAAATGGTTAGCTAAAGTTATTCACACTGTGTTA 2392
QY 302 GACTTCTCTGATCACTCCAGTGCTGACGTTGCTCTCACCGGATTCCTTACATTATCATCGCC 361
Db 2391 GATAAATTTATTGAGGCCATTATTATCAATTTTATTACAGCATTTATAACATTTTATTT 2332
QY 362 ATTGGCCCGAGCAATGCGTGGGTGGCGGATGCTCGGCACACGCTCTACAGGACTTTAT 421
Db 2331 GTAGGGCCTGTCACTCGTCAATTAGGTTATTGTTTATCTGATGGAATGACTTGGTTATAT 2272
QY 422 GATTCGSGTGGTCCAGTTCGCGGCTGTCTCTCGGTCTGGTCTACTCACCAATCGTCAATC 481
Db 2271 GAATTTGGTGGAGCTATTGTTGACTTATATTTGGTTTATATATGCGCAATCGTCATT 2212
QY 482 ACTGSGTCTGCACCAAGTCTTCCCGCCAAATTGAGCTG-----GAGCTGTTTAAAC 529
Db 2211 ACAGGAATGCACCATAGCTTTATTGCAATTGAAACGACATTAATTGCTGATGCACTAAA 2152
QY 530 CAGGTTGGATTCCTTCATCTTTGCAACGGCATCTATGGCTAATATCCGCCAGGTGGCGCA 589
Db 2151 ACAGTGGTTCAATTTATCTTCCCAATCGCAACGATGTCAAATATTCACAAGGTGGTGA 2092
QY 590 TGTTCGCGAGTGTTC---TCCTGGCGAAGAGTGAAGCTCAAGGGCCTTCAGGTTGCT 646
Db 2091 GCCTTAGCTGCATCTCTTTATCATTAAGCAAAATAAAAGGTGTGTGCTTCGCG 2032
QY 647 TCAGTGTCTCCGCTCTTCTTGGTATTACGGACCTTGCATCTTCGGTGTGAACCTTCGC 706
Db 2031 CGGGTATTTCAGCTTACTAGGAATTAAGAACACAGCAATGTTTGGTGTCAATCTTAAA 1972
QY 707 CTGCGTGGCGGTTCTTCATCCGTTATCGGTATCGGTACCGAGCTATCGTGGCGCTTTGATTGA 766
Db 1971 TTGAGATATCCATTTTATAGGTGCTGTTCAGGATCAGGTATAGGTGCGGCTTATATTTCA 1912
QY 767 CTCTTTAATATCAAGCACTTGGTTCGGCTGGCGCTGCAAGTTTCTTGGGTGTGTTCTATT 826
Db 1911 TTCTTCAAAGTAAAGCGATAGCGCTTGGTACAGCTGGATTACCTGGATTATATCTATA 1852
QY 827 GATGCTCCAGA 837
Db 1851 AATCCTACACA 1841
RESULT 9
ABK73538
ID ABK73538 standard; DNA; 975 BP.
XX AC ABK73538;
XX DT 13-AUG-2002 (first entry)
XX DE Bacillus licheniformis genomic sequence tag (GST) #829.
XX KW Differential gene expression; genomic sequenced tag; GST;
XX KW altered culture condition; environmental stress;
XX KW physiological provocation; ds.
XX OS Bacillus licheniformis.
XX PN WO200229113-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-US031437.
```





```
QY 578 CAGGTCGCGCATGTTGGCAGTGTCTTCTCTGGCGAAGTGAAGAGCTCAAGGCGCTT 637
Db |||||
QY 601 CAGGTCGCGCATGTTGGCAGTGTCTTCTCTGGCGAAGTGAAGAGCTCAAGGCGCTG 660
Db |||||
QY 638 GCAGGTCGCTTCAGGTCGTCTCGCTGTTCTTGGTATTAACGAGCTTCGAGTTCGCGTGTG 697
Db |||||
QY 661 GCTTCTTCTGCTGATTAACAGCACTTCTTGGTATTAACGAGCTTCGAGTTCGCGTGTG 720
Db |||||
QY 698 AACCTTCGCTGCGCTGCGCTGCTTCTCATCGGTATCGGTACCGAGCTATCGGTGCGCT 757
Db |||||
QY 721 AACCTTAAATACCGCTTCTCGCTTCTCATCGGTATCGGTACCGAGCTATCGGTGCGCG 780
Db |||||
QY 758 TTGATTCGCTTCTTAAATACAGCGCTTCGCTTGGCGCTGCGAGTTCCTTGGGCTT 817
Db |||||
QY 781 ATAACAGCTCTACTTAAAGTGTGCTGTATCATCGTTCAGTGGATTTCTTGGGTT 840
Db |||||
QY 818 GTTCTTATGATGCTCCAGATATGCTTCTTGGTATTAACGAGCTTCGAGTTCGCGTGTG 877
Db |||||
QY 841 CTTTCAATCAACGCAACTTCTATCCCTTCTTGGTATTAACGAGCTTCGAGTTCGCGTGTG 900
Db |||||
QY 878 ATCGATTCGCGCGAGCGATTCGCTTATGCGCTTCTTCTGCGCAACGCGAGCAT 937
Db |||||
QY 901 GTTACTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
Db |||||
QY 938 GATCCAGATCAACCGCTGCTCCAGTCCGCTGAGGACGACCAAGCGAGCAAGCA 997
Db |||||
QY 961 GAAGCAATTTGCTGAACAACATCATGTTGATCTTCCAGAAATAATACAAATCAATAGCA 1020
Db |||||
QY 998 CCGCAGAAATTTTCAAGATTCACCATCATCCAGGACCTTTGACCGGTGAGCTATT 1057
Db |||||
QY 1021 AATGCTGATGAACCTTACACAGTAGAAGACGATTTGATGCTTCTTGGTGGTGAATAACT 1080
Db |||||
QY 1058 GCACTGAGCAGCTGAGGATGCTTCTTGGCAGGAAAGCTTGGCTCGGCGCTTCC 1117
Db |||||
QY 1081 ACTCTTGTGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
Db |||||
QY 1118 ATCGTCCCAACCAAGGCGAGTGTAGTCTTCTCGGTGAGTGAAGATTTGTTGCGCATTC 1177
Db |||||
QY 1141 ATTAACCAATGGAACACTATATATCTCCAGTAGATGTTGTTCCAGTCTGATTT 1200
Db |||||
QY 1178 COATCTGCGCAATGCTTTCGAGTTCGCAACCAAGGCTGAGGATGTTCCAAATGATATC 1237
Db |||||
QY 1201 GAACTGGCGACGCTTATGACCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1248
Db |||||
QY 1238 TTGATGCAATGCTTTCGACACAGTAACTTCAACGCGACGCACTTTTAAACCGCTGAAG 1297
Db |||||
QY 1249 TTAATTCATGTTGGTATTCACACAGTATCAATGATGTTAAAGGATTCATTAAGGAT 1308
Db |||||
QY 1298 AAGCAGGCGATGAAGTCAAGCAGGAGCTGCTGTGTAATTCGATATTCGATTCGATTC 1357
Db |||||
QY 1309 GGTGCTAAACAAAGTAAAGGAGAGGTTCTTGGAACTTTGACAGCACAGTAAT 1368
Db |||||
QY 1358 AAGGCTGACAGTTATGAGTAACCGCGGATTTGTTTTCGAAATACAA 1407
Db |||||
QY 1369 ACAAATTCAGGCTTATGATGATCAACATGTTTATTTGTAACAACTCTAA 1418
Db |||||
```

## RESULT 12

AA12968

ID AAX12968 standard; DNA; 5840 BP.

XX

AC

XX

DT

19-MAR-1999 (first entry)

XX

DE

XX

KW

KW

XX

OS

XX

PN

Enterococcus faecalis genome contig SEQ ID NO:31.

Enterococcus faecalis; contig; detection; Enterococcal infection;

vaccine; attenuation; computer readable medium; ds.

Enterococcus faecalis.

WO980555-A2.

```
XX 12-NOV-1998.
PD 04-MAY-1998; 98WO-US008985.
XX 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-0046655P.
PR 14-NOV-1997; 97US-0066009P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Dillon PJ, Barash SC;
PI WPI; 1999-045171/04.
XX New isolated Enterococcus faecalis polynucleotides and polypeptides -
PT used to develop products for the detection of Enterococcus and for use in
PT vaccines for prevention or attenuation of Enterococcus infection.
XX Claim 1; Page 389-392; 2084pp; English.
XX A computer readable medium has been developed which has recorded on it
XX 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC AAX12938 to AAX13919 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence of
CC Enterococcus faecalis in samples. They can also be used for diagnosing
CC Enterococcal infection in an animal and monitoring progression of
CC disease, and for identifying agents which can be used to modulate the
CC growth or pathogenicity of Enterococcus faecalis, or another related
CC organism, in vivo or in vitro. In particular the polypeptides encoded by
CC the Enterococcus faecalis nucleotide sequences can be used in vaccines to
CC prevent or attenuate an Enterococcal infection.
XX Sequence 5840 BP; 1848 A; 944 C; 1256 G; 1779 T; 0 U; 13 Other;
```

```
Query Match 9.9%; Score 151.2; DB 2; Length 5840;
Best Local Similarity 48.5%; Pred. No. 1.4e-33;
Matches 525; Conservative 0; Mismatches 533; Indels 24; Gaps 3;
QY 329 TTGCTGCTCAGCGATTCCTTACATTCATCCCATTCGCCAGCAATCGCTGGGTGGCG 388
Db |||||
QY 1850 TTCTCTATTATGGCACTGTTACCTTCTTAGCTATTTGGTCTTATCGGACGGGTATCGGC 1909
Db |||||
QY 389 GATGCTGGCACACGCTCTACAGGACTTTATGATTTTCGGTGTCCAGTCGGCGTCTG 448
Db |||||
QY 1910 GATGCTGGGACAGGATACAAACCAATTTAGCTTTTAGTCCAATTATCGAGGTTA 1969
Db |||||
QY 449 CTCTTCGGTCTGCTTACTTCAACCAATCGTCACTGCTGTCACACAGTCTTCCGCGCA 508
Db |||||
QY 1970 TTAATGGGTTCTGTTGGCAAGTCTTGGTAATGTTTGGTATGATTTGGGGCTTTGTACCA 2029
Db |||||
QY 509 ATTGAGCTGGAGCTGTTTAAACAGGCTGGATCTTTCATCTTCGCAACGCACTATGGCT 568
Db |||||
QY 2030 ATTATGATGTTAAACTTAAACAGAGTGGCGATAGATGTTACCGATGTTATTCAGGC 2089
Db |||||
QY 569 AATATCGCCAGGCTGGCGCATGTTTGGCAGTGTCTTCTTCGCGAAGAGTGAAGCTC 628
Db |||||
QY 2090 GTTATGACAGAGCGGGGCTGCTTTAGCTGCTTTTCTTAAACAAATAATGTGAACATA 2149
Db |||||
QY 629 AAGGCGCTTGGAGGCTTTCAGGCTGCTCCGCTGTTCTTGGTATTTACGAGCTCGGATC 688
Db |||||
QY 2150 AAAGGTTGGCTTGTCTTCAAGTATTACGACTATTTTGGAAATTAACGAACTGTA 2209
Db |||||
QY 689 TTGCTGTGAACCTTCGCTGCGCTGCGCTGCTTCTTCATCGGTATCGGTACCGCAGCTATC 748
Db |||||
QY 2210 TATGCGGTGACTTTTACCATTAAGAAACCAATTTATGAGCTTGTATTTGGTGGGTATC 2269
Db |||||
QY 749 GGTGGCGCTTTGATGCACTCTTTTAAATCAAGGAGTTCGCTTGGGCGCTGAGGTTTC 808
Db |||||
QY 2270 GGTGGTGCAATTTGGCTATGAATCACGTGAAACCTTTACGTTTGGCTTGGTATG 2329
Db |||||
```

```
QY 809 TTGGGTGTTGTTCTTATTTGATGCTCCAGATATGGTCAATGTTCTTGGTGTGTCAGTTGTT 868
Db |||||
2330 TTGAGCTTGCGTGGCTTTATTTCTGCGAGACAAAAGATAGTACCGCATGATTAATCTGGT 2389
QY 869 ACCTTCTTCATCGATTTCGCGCAGCGATGCTTATGGCTTTACTTGGTTCGCGCAAC 928
Db |||||
2390 GCATTTGGTCCGGAATTCCTTATCAATGCGTTTGTCTTAACGGTTGCTTACGTT-- 2447
QY 929 GGCAGCATTTGATCCAGATGCAACCGTGTCTCCAGTGCCTCAGGAACGACCAAGCGGAA 988
Db |||||
2448 -----TTGAGATCAACCTAATCCAGAAACAGCAACTGAAAACAGAAACTGATAAG 2500
QY 989 GCGAAGACCCCGCAGATTTTCAAGATTCACCATATCCAGGCA---CCTTTGACC 1045
Db |||||
2501 ATGGTGGCACCTGTAAAAACGAATCAAGAAGACAAAATATTTTAGCAAGTCCACTTCAA 2560
QY 1046 GTTGAAGCTATTGCACTGAGCAGCGTCAGCGATGCCATGTTGCCAGCGGAAAGCTTGGC 1105
Db |||||
2561 GTTGAATTTTACCGTAGAATAAGTACAAGACCCCTGTTTTCCTTCAGGTGCTTTTAGGA 2620
QY 1106 TCGGGCGTTGCCATGCTCCCAACCAAGGGGCGAGTTAGTTTCTCCGTTGAGTGAAGATT 1165
Db |||||
2621 AAAGGTGTTGCAATTGAGCGGACTGAGGCAAACTGTATGCCACCCGAGATGTGAAATC 2680
QY 1166 GTGTGGCATTCCTATCTGGCCATGCTTTTCGCAAGTTTCGCAAGGCTGAGGATGTTCC 1225
Db |||||
2681 ACCACATTTATTTCCGACAGGACATGCTGTTGGCTTGACGACACAGAGGCG----- 2731
QY 1226 AATGTGATATCTTGATGCACATTTGTTTCGACACAGTAAACCTCAACGCGACGACTTT 1285
Db |||||
2732 ---GTTGAATTTATGATGATTTGCGATGATGATGATGATGATGATGATGATGATGATGAT 2788
QY 1286 AACCCTGCTGAAGACAGCGGCGATGAAGTCAAGCAGGGGAGCTGCTGTGTGAATTCGAT 1345
Db |||||
2789 GAATTTATCAGTGAACCAAGTGTATCTGTTAAAGAGGAGATTTGCTAGTACTTTTAT 2848
QY 1346 ATGTGATGCAATTAAGCTGAGGTTATGAGTTAAACACGCGGATTTGTTTCGAATTAC 1405
Db |||||
2849 ATTGCTGCCATTTAAAGAGCTGTTTATCGGTAGTTTACACCGATTTGTTGTAACGATACG 2908
QY 1406 AA 1407
Db 2909 AA 2910
```

## RESULT 13

```
ABS98763
ID ABS98763 standard; DNA; 5840 BP.
XX ABS98763;
AC ABS98763;
DT 18-DEC-2002 (first entry)
XX Enterococcus faecalis contig sequence #31.
DE
XX Computer readable medium; Enterococcus faecalis; microbe; growth;
KW pathogenicity; vaccine; resistance; Enterococcal infection; commercial;
KW therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;
KW biochip technology; antibacterial; modulator of nucleic acid expression;
KW contig; ds.
XX
OS Enterococcus faecalis.
XX
XX US2002120116-A1.
XX
XX 29-AUG-2002.
XX
XX 04-MAY-1998; 98US-00070927.
XX
XX 04-MAY-1998; 98US-00070927.
XX
XX (KUNSCH C A.
XX (DILLON P J.
XX
```

```
PA (BARA/) BARASH S.
XX Kunsch CA, Dillon PJ, Barash S;
XX WPI; 2002-750065/81.
XX
XX Computer readable medium having recorded on it a Enterococcus faecalis
XX nucleotide sequence useful for detecting diseases related to Enterococcus
XX infections in animals.
XX
XX Claim 1; Page; 119pp; English.
XX
XX The present invention relates to a new computer readable medium with an
XX Enterococcus faecalis nucleotide sequence. The invention is useful to
XX diagnose the presence of E.faecalis in a sample or determining the
XX presence of a specific microbe in a sample. The invention is also useful
XX for modulating the growth or pathogenicity of E.faecalis, in a vaccine to
XX confer resistance to Enterococcal infection, for commercial, therapeutic
XX and industrial purposes, and for fermenting a particular sugar source or
XX to produce a particular metabolite. The invention is useful for detecting
XX diseases related to Enterococcus infections in animals, and for detecting
XX E.faecalis using biochip technology. The present nucleic acid sequence
XX represents an Enterococcus faecalis contig DNA sequence of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at http.sequedata.uspto.gov
```

Sequence 5840 BP; 1848 A; 944 C; 1256 G; 1779 T; 0 U; 13 Other;

```
Query Match 9.9%; Score 151.2; DB 6; Length 5840;
Best Local Similarity 48.5%; Pred. No. 1.4e-33;
Matches 525; Conservative 0; Mismatches 533; Indels 24; Gaps 3;

QY 329 TTGCTGCTCAGCGATTCCTTACATTCATCGCCATTCGCGCCAGCAATGCGTGGTGGCG 388
Db |||||
1850 TTCTCTTATATGCGACCTGTTACCTTCTTAGCTTATTTGTTCTTATCGGACGGTCACTGGC 1909
QY 389 GATGTGCTGGCACAGGCTCTACAGGACTTTATGATTTTCGTTGGTCCACTGCGCGCTG 448
Db |||||
1910 GATGTGTTGGCAAGGATACAAACGCAATTTAGCTTTTAGTCCAAATTAACGAGGTTA 1969
QY 449 CTCCTTGGTCTGCTCTACTCAACCAATCGTCATCATCTGCTCGCACCGATCTCCCGCA 508
Db |||||
1970 TTAATGGGTTGCTTGTGGCAAGTCTTGGTAATGTTGGTATGATGATGATGATGATGATGAT 2029
QY 509 ATTGAGCTGGAGCTGTTTAAACAGGGTGGATCCTTCACTTCGCAACGCACTCTATGGCT 568
Db |||||
2030 ATTATGATGTTAACTTAAACAAGGTGGCGATACGATGGTACCGATGTTATTACCGCC 2089
QY 569 AATATCGCCCGGCGCGCATGTTTGGCAGTGTCTTCTCGCGAGAGTGAAGAGCTC 628
Db |||||
2090 GTTATGACAAAGCGGGGCTGCTTAGCTGTCTTTTCTTAAACAAAATGTGAACATA 2149
QY 629 AAGGCGCTTGAGGTGCTTCCAGTCTCTCCGCTGTTCTTGGTATTAACGAGGCTCGGATC 688
Db |||||
2150 AAAGGTTTGGCTTTGCTTCAAGTATTAAGACTATTTTGGATTTACTGAACCAACTGTA 2209
QY 689 TTCGGTGTGAACCTTCGCGCTGGCGTTCCTTCACTGCGTATCGGTACCGAGCTATC 748
Db |||||
2210 TATGCGTGCATTTTACCATTTGAAAAACCAATTTATTCAGCTTGTATTGTTGGTGGGTATC 2269
QY 749 GGTGGCGCTTTGATTTGCACTCTTTAATATCAAGGAGTGTGGTGGGCGCTGCAAGTTTC 808
Db |||||
2270 GGTGGTGCATTTGCGCTTATGAATCACGTTGAAAAACTTTTACGTTTGGCTTGGTTATG 2329
QY 809 TTGGGTGTTGTTTCTTATTTGATGCTCCAGATATGGTCAATGTTCTTGGTGTGTCAGTTGTT 868
Db |||||
2330 TTGAGCTTGCCTGGCTTTTATTCCTGCAGACAAAGATACTGCAACCGATGATGATGAT 2389
QY 869 ACCTTCTTCATGCGATTCGCGCAGCGATTCCTTATGCGCTTTTACTTGGTTCGCGCAAC 928
Db |||||
2390 GCAATTGGTGGCGGAATTGCCTTTATCATTCGCTTGTCTTAAACGTTTGTCTTACGTT-- 2447
```



Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

Klebsiella pneumoniae.

US6610836-B1.

26-AUG-2003.

27-JAN-2000; 2000US-00489039.

29-JAN-1999; 99US-0117747P.

(GENO-) GENOME THERAPEUTICS CORP.

Breton GL, Osborne M;

WPI: 2003-895346/82.

P-PSDB; AB066501.

New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.

Disclosure; SEQ ID NO 5847; 932pp; English.

The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This sequence encodes a Klebsiella pneumoniae polypeptide of the invention

Sequence 1452 BP; 256 A; 452 C; 429 G; 315 T; 0 U; 0 Other;

Query Match 9.4%; Score 143.6; DB 11; Length 1452;

Best Local Similarity 51.3%; Pred. No. 1.2e-31;

Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;

QY	183	TGTTTGGTTTAGATGTGCCAAGCGGTTACAGGGCACCGTGCTTCTGTGTGGTGG 242
Db	698	TCTTCGGCATCGAAGTGGCGATGATCGGCTACAGGGCACCGTCTTCCGGTGTCTGG 757
QY	243	TTTCTTGGATTCTGGCAACGATCGAGAAGTTCTCTGCACAAGCGACTCAAGGGCACTGCAG 302
Db	758	CGGTGTGGTTTATGACATGGTTCGAGAAGCGGTGGCGCGTGATCCCTGACGCGCTGG 817
QY	303	ACTTCTGATCACTCCAGTGTGACGTGCTGCTCTCACCGGATTCCTTACATTCATCGCCA 362
Db	818	ACCTGATCTCACTCCGTTCCCTGACGGTGATTTATCTCCGGCTTTATCGCCCTGCTGCTGA 877
QY	363	TTGGCCCAAGCAATGGCTGGGTGGCGATGTGTGGCACACGGTCTACAGGGACTTTATG 422
Db	878	TGGGCCCGCGCGTGGCGGCTCGCGGACGGCATTTCTGTTTATCTCAGCAGCGTTATCA 937
QY	423	ATTTCGGTGGTCCAGTCGGCGGTCTGCTCTTCGGTCTGCTTACTCACCAATCGTCATCA 482
Db	938	GCCACGCGCGGTGGCTGGGGGCGCTGCTCTTCGGGCGCTCTATTTCGGTGTGCTGTTATTA 997
QY	483	CTGGTCTGACCAAGTCTTCCCGCAATTAGCTGAGCTGTTTAAACCA-----GG 533
Db	998	CCGGTATCCATCACAGTTCATGCGCATCGAGCGCGGACTGCTGGGCAACCCATCGATTG 1057
QY	534	GTGGATCTCTTCACTTCGCAAGCGCATCTATGGCTAATATCGCCAGGTCGGCATGTT 593
Db	1058	GGGTCAACTTCTCTGTCGGATCTGGGCGATGGCCAAAGCTCGCCAGGCGCGCTGCT 1117
QY	594	TGGCAGTGTCTTCTTGGCGAAGATGAAGAACTCAAGGGCCCTTCAGGTGCTTCAGGTG 653
Db	1118	TTGCGGTGTGGTTTAAACCAAGATGCCAAATAAAGCTATACCTTCGCGTGGCGT 1177
QY	654	TCTCCGCTGTTCTTGGTATTCAGGACCTCGCATCTTCTGGTGTGAACCTTCGCTGGCGT 713
Db	1178	TTTCGGCGATGTGGGGATCACCGAGGGGCAATCTTCGGGATTAACTTCGCTTTGTGA 1237

QY	714	GGCGTTTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATGCACTCTTAA 773
Db	1238	AACCGTTTCATCGCGCGCTGGTGGCGGTGCCCGCGCGCGCTGGGTGGTGTTCGATGC 1297
QY	774	ATATCAAGGCAGTTGGCTTGGCGCTGCGAGTTTCTTGGGTGTGTTTCTATTGATGCTC 833
Db	1298	ACGTCTACATGACCGCGGTGGGCTGACCGCGATCCCGGGAATGGCTATCGTCAGGCCA 1357
QY	834	CAGATATGTCATGTTCTTGGTGTGTCAGTTGTACCTTCTTCATCGCATTCG 887
Db	1358	GTCGCTGCTGAACTACATTATCGGAATGCGCATCGCTTCGCGGTGGCTTCG 1411

Search completed: October 30, 2004, 14:04:57

Job time : 734 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 13:46:16 ; Search time 154 Seconds  
(without alignments)  
7047.887 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 cccatggcatcgcgcgtt.....gttgaaaccttgatgttgcg 1527

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/1/ina/6H COMB.seq:\*

5: /cgn2\_6/prodata/1/ina/PTUS COMB.seq:\*

6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	190.6	12.5	2913	4	US-09-710-279-3464, Ap
C 2	171.6	11.2	1491	4	US-09-134-000C-2417, Ap
C 3	143.6	9.4	1386	4	US-09-489-039A-6203, Ap
C 4	143.6	9.4	1452	4	US-09-489-039A-5847, Ap
C 5	137.4	9.0	732	3	US-09-134-001C-1277, Ap
C 6	126.6	8.3	1368	4	US-09-583-110-297, Ap
C 7	126.6	8.3	3895	4	US-08-961-527-201, Ap
C 8	117.6	7.7	1884	4	US-09-583-110-463, Ap
C 9	114.6	7.5	951	4	US-09-134-000C-2077, Ap
C 10	109.6	7.2	3615	1	US-08-920-812-17, Ap
C 11	109.6	7.2	3615	1	US-08-920-827-17, Ap
C 12	109.6	7.2	3615	1	US-08-921-177-17, Ap
C 13	109.6	7.2	3615	1	US-08-362-577C-17, Ap
C 14	109.6	7.2	3615	2	US-08-920-828-17, Ap
C 15	109.4	7.2	1905	4	US-09-543-681A-2847, Ap
C 16	107.8	7.1	9769	4	US-08-961-527-30, Ap
C 17	97.8	6.4	465	2	US-08-673-190A-3, Ap
C 18	97.2	6.4	1971	4	US-09-489-039A-1041, Ap
C 19	90	5.9	1428	4	US-09-489-039A-1806, Ap
C 20	86.2	5.6	30246	4	US-08-956-171E-56, Ap
C 21	86.2	5.6	30246	4	US-08-781-986A-56, Ap
C 22	83.2	5.4	1284	4	US-09-107-532A-1876, Ap
C 23	81	5.3	1887	4	US-09-107-532A-1634, Ap
C 24	76.4	5.0	465	4	US-08-956-171E-1317, Ap
C 25	76.4	5.0	465	4	US-08-781-986A-1317, Ap
C 26	71.6	4.7	357	2	US-08-673-190A-6, Ap
C 27	71	4.6	270	4	US-09-134-000C-2419, Ap

28	70.4	4.6	2013	4	US-09-134-000C-1201, Ap
29	67	4.4	2181	4	US-09-583-110-121, Ap
30	67	4.4	8494	4	US-08-961-527-163, Ap
31	64.4	4.2	1446	4	US-09-543-681A-984, Ap
32	60.4	4.0	1437	3	US-09-134-001C-2228, Ap
33	60.4	4.0	2550	4	US-09-710-279-4343, Ap
34	58.8	3.9	1752	4	US-09-107-532A-3157, Ap
35	57.6	3.8	1971	4	US-09-107-532A-1429, Ap
36	56.8	3.7	29555	4	US-08-956-171E-206, Ap
37	56.8	3.7	29555	4	US-08-781-986A-206, Ap
38	55.8	3.7	1846	4	US-09-634-238-183, Ap
39	55.4	3.6	1896	4	US-09-107-532A-248, Ap
40	54.6	3.6	591	4	US-09-543-681A-3487, Ap
41	53.2	3.5	2996	4	US-09-710-279-3809, Ap
42	53.2	3.5	3081	4	US-09-710-279-4310, Ap
43	53.2	3.5	3932	4	US-09-710-279-3420, Ap
44	52.2	3.4	315	4	US-09-134-000C-2078, Ap
45	51.4	3.4	1839	4	US-09-583-110-250, Ap

## ALIGNMENTS

## RESULT 1

US-09-710-279-3464/c  
; Sequence 3464, Application US/09710279  
; Patent No. 6703492

GENERAL INFORMATION:

APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PUB480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 3464  
LENGTH: 2913  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-3464

Query Match	12.5%;	Score	190.6;	DB	4;	Length	2913;
Best Local Similarity	55.0%;	Pred. No.	3.7e-48;				
Matches	468;	Conservative	0;	Mismatches	364;	Indels	19;
Gaps	4;						
QY	5	TGGCATCTGCGCGTTCGGTTCCTTGGCGCGCGTATTTGGTATCGCATGTTTCCCGAGCTTG	124				
Db	2690	TTGCAATGACCTTTACATTTATACCAATCTTATTGGTTTATGTCAGCTAAGCAT	2631				
QY	65	TGGCGGCGCAATGAGTTCCTTGGCGCGCGTATTTGGTATCGCATGTTTCCCGAGCTTG	124				
Db	2630	TTGGTGGTAATCTTATTATTAGTGAGC-TCTAGGTATGATCTTGTTCATCTGATTTG	2572				
QY	125	GTGAACGGCTACGACGTGGCGCCACCATGGTGGCGGCGAA---ATGCCAATGTGGTCC	181				
Db	2571	ATGAGTGCATATGATTTTCCAAAAGCTTTAGAAGAAGAAAAGCTATTCACACATGGAT	2512				
QY	182	CTGTTTGGTTAGATTTGCCCAACCGGTTCACAGGGCACCGTCTCTCTGTGCTGGTG	241				
Db	2511	GTCTTTGGCTACATATTAAATGAAGTAGTTATCAGGGACAGTATTACCTATGCTGTA	2452				
QY	242	GTTCCTTGGATTCTGGCAACGATCGAGAAGTTCCTGTCACACCGACTCAAGGGCACTGA	301				
Db	2451	GCAATATATTTTACGTACGATTGAAAATAGTTACGTAAAGTTATTCCAACTGTGTTA	2392				
QY	302	GACTTCCTGATCACCCTCCAGTGTGCTGCTGCTCAGCGGATTCCTTACATTCATCGCC	361				
Db	2391	GATAATTTTTCAGCCCATTTATCAATTTTATTACAGCATTTATACATTTTATTT	2332				

QY	362	ATTGGCCCGCAATCGCTGGGTGGCGATGTCTGGCACAAGGTCTACAGGACATTTAT	421
DB	2331	GTAGGGCTGTCACTCGTCAATTAGGTTATTTGGTTATCTGATGGAATGACTTGGTATAT	2272
QY	422	GAATTCGGTGGTCCAGTCGGCGGTCTGCTCTTCGGTCTGTCTACTCACCAAATCGTCATC	481
DB	2271	GAATTTGGTGGAGCTATTGTTGGACATATATTTGGTTTATATATGCGCCAATCGTCATT	2212
QY	482	ACTGGTCTGCACAGATCCTTCCCGCCAAATGAGCTG-----GAGCTGTTTAAAC	529
DB	2211	ACAGGAATGCACCATAGCTTTATTGCAAGTTGAAACGACATTAATTCGTGATCGCACTAAA	2152
QY	530	CAGGGTGATCCCTTCATCTTCGCAAGGCACTATGCTGTAATATCGCCAGGTCGGCA	589
DB	2151	ACAGGTGGTCAATTTATCTTCCCAATCGCAACGATCAAAATATGCAAGGTGGTGA	2092
QY	590	TGTTGGCAGTGTCT---TCCTGGCGAAGAGTGAAGCTCAAGGCGCTTCAGATGCT	646
DB	2091	GCITTAGCTGATCTTTATCATTTAAGCAAAATAAAAAATTAAGGTGTGCTCCGCG	2032
QY	647	TCAGGTCTCCCGTGTCTTGTGTAATTAACGAGCTCGCATCTTCGGTGTGAACCTTCGC	706
DB	2031	CGCGGTATTTACGCTTTTACTAGGAATTTACAGAACAGCAATGTTTGGTGTCAATCTTAA	1972
QY	707	CTCGCTGGCGGTCTTTCATCGGTATCGTACCGCACTATCGTGTGCGCTTTGATTGCA	766
DB	1971	TTGAGATATCCATTTATAGTGTCTTTGACGATCAGTATAGTGGCGCTTATATTCA	1912
QY	767	CTCTTTAATATCAAGGCAGTTGCTTGGGCGCTGCAGTTTCTTGGGTGTGTTCTTATT	826
DB	1911	TTCTTCAAGTAAAGCGATAGCGCTTGGTACAGTCGATTACCTGGATTATATCTATA	1852
QY	827	GATGCTCAGA	837
DB	1851	AATCCTACACA	1841
RESULT 2			
US-09-134-000C-2417			
; Sequence 2417, Application US/09134000C			
; Patent No. 6617156			
; GENERAL INFORMATION:			
; APPLICANT: Lynn Doucette-Stamm et al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO			
; FILE REFERENCE: 032796-032			
; CURRENT APPLICATION NUMBER: US/09/134,000C			
; PRIOR FILING DATE: 1998-08-13			
; PRIOR FILING DATE: 1997-08-15			
; NUMBER OF SEQ ID NOS: 6842			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 2417			
; LENGTH: 1491			
; TYPE: DNA			
; ORGANISM: Enterococcus faecalis			
US-09-134-000C-2417			
Query Match 11.2%; Score 171.6; DB 4; Length 1491;			
Best Local Similarity 47.7%; Pred. No. 1.8e-42;			
Matches 596; Conservative 0; Mismatches 624; Indels 30; Gaps 2;			
QY	176	TGTCCTCCCTGTTGGTTTATAGTTTGCACAGCGGTACAGGCGACCGTCTTCCTGTG	235
DB	181	TGGATATATTTTGGATACCATGTTGCACAAAACAACTATGCTACCAAGTAATTCGGTA	240
QY	236	CTGGTGGTCTCTTGATTTCTGCAACGATCGAGAAGTTCTTCGCAACGCACTCAAGGC	295
DB	241	TTAGCTTCGGTATATCTTTTGTCAATATTTGAAAATATTTTCATAAAAACCTTCCTCA	300
QY	286	ACTGCAGACTTCCTGATCACTCCAGTGTGAGCTTGTGCTCACGGATTCCTTCAATTC	355

## RESULT 3

US-09-489-039A-6203/c  
; Sequence 6203, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 6203  
; LENGTH: 1386  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-6203

Query Match 9.4%; Score 143.6; DB 4; Length 1386;  
Best Local Similarity 51.3%; Pred. No. 7.7e-34;  
Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;  
QY 183 TGTGTTGGTTTAGATGTTGCCAAGCGGTTACAGGGCACCGTGCTTCTGCTGCTGGTGG 242  
Db 794 TCTTCGGCATCGAAGTGGCGATGATCGGCTACAGGGCACCGTCTTCCGGTGGTGGTGG 735  
QY 243 TTTCTTGGATTCTGGCAAGATCGAGAAGTTCTCTGCACAAGCGACTCAAGGGCACCTGCAG 302  
Db 734 CGGTGTGGTTTATGAGCATGTTGCGAAGCGGCTGCGCGCGTGAATCCCTGACGCGCTGG 675  
QY 303 ACTTCTGATCACTCAGTGTCTGCTGCTGCTCACCAGATTCCTTACATTCATCGCCA 362  
Db 674 ACCTGATCTCACTCCGTTCTCTGACGGTGATTAATCTCCGGCTTTATCGCCCTGCTGTA 615  
QY 363 TTGGCCCAAGCAATGCGTGGTGGGCGATGTGCTGGCACACGGTCTACAGGACCTTATG 422  
Db 614 TCGGCCCGCGGTGCGCGCTCGGGGACGGATTTGTTTATCTCAGCAGCTTATCA 555  
QY 423 ATTTGCGTGGTCAAGTGGCGGCTGCTCTTCGGTCTGCTTACTACCAATCGTCATCA 482  
Db 554 GCCACGCGGCTGGCTGGCGGCGCTGCTGTTCCGGCGGCTTATTCGGTGAATCGTTATTA 495  
QY 483 CTGCTGTCACACAGTCTTCCCGCAATGAGCTGAGCTGAGCTGTTAAACA -----GG 533  
Db 494 CCGGTATCCATCACAGCTTCCATGTCATCGAGGCGGAGCTGCTGGGCAACCCATCGATTG 435  
QY 534 GTGGATCTTCTATCTTCGCAAGGCAATCTATGGCTAATATCCCGAGGCTGGCGATGTT 593  
Db 434 CGGTCAACTTCTGCTGCGGATCTGGGCGATGCGCAACGTGCGCCAGGGCGGCGCTGCT 375  
QY 594 TGGCAGTGTCTTCTGCGGGAAGATGAAAGCTCAAGGCGCTTGCAGGTGCTTCAGGTG 653  
Db 374 TTGCGGTGTGTTTAAACCAAGATGCCAAATAAAGCTATACCTCGCGCTGGCGT 315  
QY 654 TCTCGCTGTTCTTGGTATACGGAGCTTCGATCTTCGGTGTGAACCTTGCCTGCGCT 713  
Db 314 TTTCCGCGATGCTGGGATCACCGAGGCGGCAATCTTCGGGATTAACCTGCGCTTGTGA 255  
QY 714 GCGCGTCTTCTATCGGTATCGGTACCGAGCTATCGGTGGCGCTTGTATGACATCTTTA 773  
Db 254 AACCGTTCATCGCGGCTGTTGGGCGGTGCGCGCGGCGGCGCTTGGTGGTGGTGGTGC 195  
QY 774 ATATCAAGCGAGTTCGGTGGCGCTGAGGTTCTTGGTGTGTTTCTTATGATGCTC 833  
Db 194 ACGTCTACATGACCGGCTGGGCTGACCGCGATCCCGGAATGCTATCGTGCAAGGCA 135  
QY 834 CAGATATGCTATGTTCTGTTGTTGTCGAGTTGTTTACTTCTTATCGCATTCG 887  
Db 134 GCTCGCTGCTGAATACATTAATCGGAATGGGATCGCTTCCGCGTGGCCTTCG 81

## RESULT 4

US-09-489-039A-5847  
; Sequence 5847, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5847  
; LENGTH: 1452  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5847

Query Match 9.4%; Score 143.6; DB 4; Length 1452;  
Best Local Similarity 51.3%; Pred. No. 7.9e-34;  
Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;  
QY 183 TGTGTTGGTTTAGATGTTGCCAAGCGGTTACAGGGCACCGTGTCTGCTGCTGGTGG 242  
Db 698 TCTTCGGCATCGAAGTGGCGATGATCGGCTACAGGGCACCGTCTTCCGGTGGTGGTGG 757  
QY 243 TTTCTTGGATTCTGGCAAGATCGAGAAGTTCTCTGCACAAGCGACTCAAGGGCACCTGCAG 302  
Db 758 CGGTGTGGTTTATGAGCATGTTGCGAAGCGGCTGCGCGCGTGAATCCCTGACGCGCTGG 817  
QY 303 ACTTCTGATCACTCAGTGTCTGCTGCTCACCAGATTCCTTACATTCATCGCCA 362  
Db 818 ACCTGATCTCACTCCGTTCTCTGACGGTGATTAATCTCCGGCTTTATCGCCCTGCTGTA 877  
QY 363 TTGGCCCAAGCAATGCGTGGTGGGCGATGTGCTGGCACACGGTCTACAGGACCTTATG 422  
Db 878 TCGGCCCGCGGCTCGCGCTCGCGGCGATTTGCTTTATCTCAGCAGCTTATCA 937  
QY 423 ATTTGCGTGGTCCAGTTCGGGCTGCTCTTCGGTCTGCTGCTACTACCAATCGTCATCA 482  
Db 938 GCCACGCGGCTGGCTGGCGGCTGCTGTTCCGGGCGCTTATTCGGTGAATCGTTATTA 997  
QY 483 CTGCTGTCACACAGTCTTCCCGCAATGAGCTGAGCTGAGCTGTTAAACA -----GG 533  
Db 998 CCGGTATCCATCACAGCTTCCATGTCATCGAGGCGGAGCTGCTGGGCAACCCATCGATTG 1057  
QY 534 GTGGATCTTCTATCTTCGCAAGGCAATCTATGGCTAATATCCCGAGGCTGGCGATGTT 593  
Db 1058 GCGTCAACTTCTGCTGCGGATCTGGGCGATGCGCAACGTGCGCCAGGCGGCGCTGCT 1117  
QY 594 TGGCAGTGTCTTCTGCGGGAAGATGAAAGCTCAAGGCGCTTGCAGGTGCTTCAGGTG 653  
Db 1118 TTGCGGTGTGTTTAAACCAAGATGCCAAATAAAGCTATACCTCGCGCTGGCGT 1177  
QY 654 TCTCGCTGTTCTTGGTATACGGAGCTTCGATCTTCGGTGTGAACCTTGCCTGCGCT 713  
Db 1178 TTTCCGCGATGCTGGGATCACCGAGGCGGCAATCTTCGGGATTAACCTGCGCTTGTGA 1237  
QY 714 GCGCGTCTTCTATCGGTATCGGTACCGAGCTATCGGTGGGCGCTTGTATGACTCTTTTA 773  
Db 1238 AACCGTTCATCGCGCGCTGTTGGGCGGTGCGCGCGGCGGCGCTTGGTGGTGGTGGTGC 1297  
QY 774 ATATCAAGCGAGTTCGGTGGCGCTGACAGTCTTCTGCGGTGTTTCTTATGATGCTC 833  
Db 1298 ACGTCTACATGACCGGCTGGGCTGACGGGATCCCGGAATGCTATCGTGCAAGGCA 1357  
QY 834 CAGATATGCTATGTTCTGTTGTTGTCGAGTTGTTTACTTCTTATCGCATTCG 887  
Db 1358 GCTCGCTGCTGAATACATTAATCGGAATGGGATCGCTTCCGCGTGGCCTTCG 1411

RESULT 6

US-09-583-110-297  
; Sequence 297, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lyvn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 297  
; LENGTH: 1968  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-297

Query Match 8.3%; Score 126.6; DB 4; Length 1968;  
Best Local Similarity 47.0%; Pred. No. 1.7e-28;  
Matches 581; Conservative 0; Mismatches 629; Indels 27; Gaps 5;

Qy	174	TGTGGTCCCTGTTGGTTTGAATGTTGCCAAGCCGGTTTACCAGGGCACCGTGCTTCCTG	233
Db	662	TTTCGAATTTTGGTATTTTACTGTATTCGTATTTGGTTTACCAGCCCAAGTTATCCCGAG	721
Qy	234	TGCTGGTGGTTTCTTTGGATTCTGGCAAGATCGAGAAAGTTCTCGCACAGCGACTCAAGG	293
Db	722	CTTGTCTGCAGGTTTGAGTCTGTCTTATCTTGAATTTTGGGCGCAACATATCCCGAG	781
Qy	294	GCACGTGCAGATTCCTGTATCACTCCAGTCTGCAGCTTGTCTGCACCGGATTCCTTACAT	353
Db	782	AAGTCATTTCCATGATTTTGTGTACCTTTTGTCTCAATGATTCAGCCTTGAATTTGGCTC	841
Qy	354	TCATCGCCATTGGCCCGACCAATCGCTGGTGGCGGATGCTGGGCACACGGTCTACAGG	413
Db	842	ATACTGTCTGGGGCCAAATCGGTGGACAAATGGACAGGACTTTCATCAGTTGTATTGG	901
Qy	414	GACTTT---AATGATWTCGGTGGTCCAGTGGCGGTTCTGCTCTTGGGTCTGCTCTACTAC	470
Db	902	CAGGTTTAAACGGTCTCTGTAATGGCTCTTCGGTGCATTTTGGTGCCCTCTACGCTC	961
Qy	471	CAATCGTCACTACTGGTCTGCACAGTCTTCCCGCCAAATTCAGCTGGAGCTGTTT---A	527
Db	962	CATTGTGCATCACAGGCTCGCACCAATGACCAATGCCCATTGATACACAAATGATTCGGG	1021
Qy	528	ACCAGGGTGGATCTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGGTGGCGG	587
Db	1022	ATGCTGGTGGCACTGCCCTCTGGCCTATGATTTGCTCTTCTAATATTGCTCAAGGCTCAG	1081
Qy	588	CATGTTTGGCAGTGTCTTCCTGGCGAAGAGTGAAAAGCTCAAGGCC---TTGCGAGTG	644
Db	1082	CGGTGTTGCCATTATTTTCATGCATCGCCATGATGAGTGTGAGGCTCAGGTTTCACTTC	1141
Qy	645	CTTCAGGTCTCCCGCTGTTCTTTGTTATTACGAGCGCTCGATCTTCGGTGTCAACCTTC	704
Db	1142	CTGCACCAATTCAGCCTATCTCGTGTTACAGAACAGCTCTTTTGGGTTAACGTAA	1201
Qy	705	GCCTCGCTGGCCGTTCTTTCATCGGTATCGGTACCGCATATCGGTGGCGCTTGAATG	764
Db	1202	AATATATTTATCCATTTTGTGTGGATGACTGTGTTAGCCCTTCAGCGCATGTTATTCGG	1261
Qy	765	CACCTCTTAATATACAGGCAGTTGGTCTGGCGCTCAGGTTTCTTTGGGTGTTGTTCTTA	824
Db	1262	TTACTTTTTAATGTAACCTGGGCTTCATTGGTATTCGGTGTTCAGGATTAATCTCTCTA	1321
Qy	825	TTGATGCTCCAGATATGGTCATGTTCTTGGTGTGCGAGTTGTATCCTTCTTCATGCGAT	884

RESULT 5

US-09-134-001C-1277

Sequence 1277, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 1277

LENGTH: 732

TYPE: DNA

ORGANISM: Staphylococcus epidermidis

FEATURE:

NAME/KEY: unsure

LOCATION: (20),(22),(32)

OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.

US-09-134-001C-1277

Query Match 9.0%; Score 137.4; DB 3; Length 732;

Best Local Similarity 56.4%; Pred. No. 4.2e-32;

Matches 305; Conservative 0; Mismatches 221; Indels 15; Gaps 2;

Qy 312 TCACGCCAGTGTGCTGCTGCACCGGATTCCTTACATTCATCGCCATTCGCCCCAG 371

Db 80 TGACGCCATTATTCAATTTTATTACAGCAATTATAACATTTTATTGTAGGCCCTG 139

Qy 372 CAATGCGTGGGTGGCGGATGTCTGGCACACGGTCTACAGGACTTTTATGATTTCCGTTG 431

Db 140 TCACTCGTCAATTAGGTATTGTTTATCTGATGGATTCGACTTGGTTATATGAATTTGGTG 199

Qy 432 GTCCAGTCGGGGTCTGCTCTTCGGTCTGGTCTACTACCAATCGTCATCACTGGTCTGC 491

Db 200 GAGCTATTGGTGAGCTTATTTTGGTTTATTAATGCGCCAAATCGTCATACAGGAATGC 259

Qy 492 ACCAGTCCTTCGCCCAATTGAGCTG-----GAGCTGTTTAAACAGGGTGGAT 539

Db 260 ACCATATCTTTATTCAGTTGAAACAGCATTAATTCCTGATCGACTAAAACAGGTGGTT 319

Qy 540 CTTCAATCTTCGACAGGATCTATGCTTAATATCGCCAGGTGCGGCATGTTTGGCAG 599

Db 320 CATTTATCTTCCCAATCGCAAGTGTCAAAATTTGCACAAGGTGGTGAGCTTTAGCTG 379

Qy 600 TGTTCT--TCCCTGCGAAGAGTGAAGAGCTCAAGGGCCCTTCGAGGTGCTTCAGGTGCT 656

Db 380 CATCTTTATCAATTAAGCAAAATATAAATTTAAAGTGTGTTGCTCCGCGCGGTATTT 439

Qy 657 CGCTGTTCTTGGTATTAACGAGCCTTCGATCTTCGGTGTGAACCTTCGCTCGCTGGC 716

Db 440 CAGCTTTACTGAATTTACAGAACAGCAATGTTTGGTGTCAATCTTAAATTCAGATATC 499

Qy 717 CGTTCTTCATCGGTATCGTACCGGAGCTATCGGTGGCGCTTTGATTGGCACTCTTAATA 776

Db 500 CATTTATAGGTCTGTTCAGGATCAGGTATAGGTGGCGCTTATATTTCAATCTTCAAG 559

Qy 777 TCAAGGCAAGTTCGTTGGCGGTGTCAGGTTTCTTTGGGTGTGTTTCTATTGATGCTCCAG 836

Db 560 TAAAGCGATAGGCTTGTACAGCTGGATACCTGGATTATATCTATAAATCCCTACAC 619

Qy 837 A 837

Db 620 A 620

Db	1322	TTCAACTCAATACATGCTGCCATTTGCAGGAATATGCTAGTTCGCAATGTTGTTCCAA	1381
Qy	885	TCGGCGCAGCAANTGCTTATGGCCTTTACTTTGGTTTCGCCGCAACGGCAGCAATTGATCCAG	944
Db	1382	TGCTC-----TTGACTTCTCTTTTCGCAAGGCTGCTCTTCACAAAAATAGAGGGCG	1435
Qy	945	ATGCAACCGCTGCTCCAGTGCCTTCAGGACGACCAACGAAGCGAAGCAGAAGCACCOCGAG	1004
Db	1436	ATACGAACCTTGCAGCGAGAATTGCTCTCAAGAAGAAGACGAATTTGTGAGCCATGAAC	1495
Qy	1005	AAATTTTCAAAACGATTCACCATCATCCAGGCACCTTTGACCGGTGAAGCTATTGCACTGA	1064
Db	1496	CAGTAGAACTTACTTTCGGTAGAAATATCAGCCCACTAACTGGCCATAGTGAAGAATTTGA	1555
Qy	1065	GCAGCGTCAGCATGCCATGTTTGCCAGCGAAAGCTTGCTCGGGCGTTGCCATCGTCC	1124
Db	1556	GTCAAAGCGACGGATCCCTGTTTTTGATCAGGTGTCATGGGCGCAAGGTCTGGTCATTGAAC	1615
Qy	1125	CAACCAAGGGCAGTTAGTTCTCCGGTGAGTGAAGAAGATGTGTGGCATTGCCATCTG	1184
Db	1616	CAAGCAGAGTGAGTTGACCTCTCCAGTTAATGGAAACAGTGAAGGTTCTTTTCCCTACCA	1675
Qy	1185	GCCATGCTTTCCGAGTTTCGACCAAGGCTGAGATGTTCCAAATGTGGATATCTTCATGC	1244
Db	1676	AGCATGCCATCGGCATTGTCTCTGACGAGGAGTTGAAT-----TGCTCATCC	1723
Qy	1245	ACATTGGATTTGACACAGTAAACCTCAACGCGACGCACTTTAAACCGCTGAAGAAGCAGG	1304
Db	1724	ACATCGGTATGGATACAGTAGGCTCTTGATGCGAAAAGGTTTTGAAAGTCTTTGTAGTCCAAG	1783
Qy	1305	GCGATGAAGTCAAAGCAGGGGAGCTGCTGTGTGAATTCGATATTGTATGCCATTAAAGGCTG	1364
Db	1784	GAGATCAGTTATAGTTGGCCAGCACTGAATCGTTTTGATATGTAATGTAATGTAAGGCTG	1843
Qy	1365	CAGGTTATAGGTAAACACCCGAGTTGTTTCGAA	1401
Db	1844	CAGGTCGTGGTGCAGAAACCTCCTGTTATCATCAACCAA	1880

## RESULT 7

```

US-08-961-527-201/c
; Sequence 201, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

```



GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2077  
LENGTH: 951  
TYPE: DNA  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-2077

Query Match 7.5%; Score 114.6; DB 4; Length 951;  
Best Local Similarity 48.4%; Pred. No. 5.5e-25;  
Matches 429; Conservative 0; Mismatches 434; Indels 24; Gaps 3;

QY 524 TTTAACCAGGTGGATGCTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGT 583  
Db 10 TTAACACAGGTGGCATAGCTACCGATGTTATACCGCCGTTATTCACAGGC 69  
QY 584 GGGCGATGTTGGCAGTGTCTTCCTGGCGAAGAGTGAAAAGCTCAAGGGCCCTTCAGGT 643  
Db 70 GGGGCTGCTTATAGTCTCTTTCCCTAACAAAAATGTGAACTAAAGSTTGGCTTG 129  
QY 644 GCTTCAGGTGTCCTCGCTCTCTTGATATACGAGCCCTGCGATCTTCGGTGGAACCTT 703  
Db 130 TCTTCAAGTATTACGACTATTTTGGAAATTAAGTAACTGATATATGCGGTGACTTTA 189  
QY 704 CGCCTGCGTGGCGGTCTTCATCGGTATCGGTACCGCAGCTATCGTGGCGCTTTGATT 763  
Db 190 CCATTGAAAAACCAATTATTCGACGTTGATTTGGTGGCGGTATCGGTGGCAITTTG 249  
QY 764 GCATCTTTAATATCAAGCAGTGTGGTGGCGTGCAGGTTTCTTGGGTGTGTTTCT 823  
Db 250 GCTATGAATCAGTGAAGAACTTTACGTTTGGCTTGGTGTAGTATGTTGAGCTTGCCTGC 309  
QY 824 ATTGATGCTCCAGATATGTCATGTTCTTGGTGTGCGAGTTGTACCTTCTTCATGCA 883  
Db 310 TTTATTCCTGCAGACAAAGATATGCAACCGATGATTACTGTGTCGAATGCTGCGCA 369  
QY 884 TTCGGCGCAGCATGCTTATCGCCTTTACTTGGTGTTCGGCAACGGCAGCATGATCCA 943  
Db 370 ATTGCTTTATCATTTGGCTTGTCTTAACGTTGT-----CTTACGTTTGAAGAT 420  
QY 944 GATGCAACCGCTGCTCCAGTGCCTGCAGGAACGACCAAGCCGAGCAGAACCCCGCA 1003  
Db 421 CAACCTAATCCAGAAACAGCAACTGAAAAACAGAAACCTGATAAGATGTTGGCACCCTGA 480  
QY 1004 GAATTTTCAACGATTCACCATCATCAGGCA---CCTTGACCGGTGAAGCTATTGCA 1060  
Db 481 AAAACGAATCAAGAGACAAAAATATTTAGCAAGTCCACTTCAAGGTGAAATTTTACCG 540  
QY 1061 CTGACGAGCGTCAGCGATGCCATGTTTGGCAGCGGAAAGCTTGGCTCGGGCGTGCATC 1120  
Db 541 CTAGAAAAGTACAGACCCCTGTTTGTCTCAGGTGCTTTAGGAAAAGGTGTTGCAATT 600  
QY 1121 GTCCCAACCAAGGGCAGTTAGTTCTTCGGTGTAGTGAAGAAATGTTGGTGGCAATCCCA 1180  
Db 601 GAGCGGACTGAAGGCAAACTGATGACCCCGCAGATGTTGAATCACCACATTAATTCG 660  
QY 1181 TCTGCCATGCTTTCGAGTTCGACCAAGGCTGAGGATGTTCCAAATGTGGANATCTTG 1240  
Db 661 ACAGACATGCTGTTGGCTTGCAGCAACACAGAG-----GCGTTGAATTA 708  
QY 1241 ATGCATTTGTTTCGACACAGTAACCTCAACGGCAGCAGCTTTAAACCCGCTGAGAG 1300  
Db 709 ATGCATTTGCAUGATACGGTCGAATTAGATGTGTAAGGCTTTGAATTAATCAAGTGA 768

QY 1301 CAGGCGCATCAAGTCAAAGCAGGCGAGCTGCTGTGTAATTCGATTTGATGCCATTAA 1360  
Db 769 CRAAGTGATCTGTTTAAAAAGAGATTGCTAGTTACTTTTGAATTTGCTGCCATTAA 828  
QY 1361 GTCGAGGTATGAGTAACCAACCGCGATGTTGTTTCGAATTACAA 1407  
Db 829 GAAGCTGGTTATCCGGTAGTTACACCGATTGTCGTAACGAATACGAA 875

RESULT 10  
US-08-920-812-17/c  
Sequence 17, Application US/08920812  
Patent No. 5763188  
GENERAL INFORMATION:  
APPLICANT: Ohno, Tsuneya  
APPLICANT: Matsuhisa, Akio  
APPLICANT: Uehara, Hirotsugu  
APPLICANT: Eda, Soji  
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,812  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3615 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-24  
US-08-920-812-17

Query Match 7.2%; Score 109.6; DB 1; Length 3615;  
Best Local Similarity 46.3%; Pred. No. 4.4e-23;  
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;

QY 312 TCATCCAGTCTGACGTTGCTGCTCACCGGATTCCTTACATTCATCCCAATGSCCAG 371  
Db 2495 TCACACATTGCTATGCTGATGTTATCACACCGTCACCTTTCTGCTGGTGGGCGC 2436  
QY 372 CAATCGCGTGGTGGCGATGCTGSCACAGGCTCTACAGGACTTTATGATTCGGTG 431  
Db 2435 TATCAACTGTAAGCGAACTGATTGCGCGGTTATCTCTGGCTTTATCAGGCGGTT 2376  
QY 432 GTCCAGTCCGCGGTCTGCTCTTCGCTGCTGCTTACTCACCAATCGTCATCACTGCTGTC 491

Db 2375 CTGCAATTGGCGGCGGTAATGGGCGCTTCTGGCAAACTTCTCGTCAAGTTCGGAAGTGC 2316  
QY 492 ACAGTCTCTCCGCCCAATAGCTGAGCTGTTTAACACAGGTTGGATCCTTCACTTCG 551  
Db 2315 ACTGGGCGCTGGTCCGGCTGTATCAATACTTCAACCGTGTCTGGGCTAGACACCAAGA 2256  
QY 552 CAACGGCATCTATG--GCTAATATCGCCAGGTTGGGCAATGTTGGCAGTGTCTTCC 608  
Db 2255 TCCGCTGTATATCCCGCCCAATTAAGGCGAGTTCGGGCGGCGCTCGGCTCTCTCT 2196  
QY 609 TGGCGAAGAGTGAAGCTCAAGGCGCTTTCAGGTGCTTCAAGTGTCTCCGCTGTCTTG 668  
Db 2195 GCGAACCGATGCGCAGAAAAAGTGTGGCGGATCAGCGCGCTGACGAGTCTGTG 2136  
QY 669 GTATTACGGACCTGCGATCTTCGGTGAACCTTCGCTCGGCTGGCGGCTTCTTCATCG 728  
Db 2135 GTATCACGGAACCAAGCGGTATATGCGCTCAACCTCGCGCGTAAGTACCCCTTGTATCG 2076  
QY 729 GTATCGGTACCGCAGCTATCGTGGCGCTTGTGATGTCACCTCTTTAATATCAAGCAGTTG 788  
Db 2075 CCGTATCAGTGGGCTTGGGGCCACCATATTATGGCTACGCGCAACGAAGTCTACT 2016  
QY 789 CGTTGGCGCTGCAAGTTCTTGGGTGTGTTCTATTGATGCTCCAGATATGTCATGT 848  
Db 2015 CTTTGGTTTCCAAATATTTTACCTTCATGCAAAACCATCCCGTCAACGGGAATTGATT 1956  
QY 849 TCTTGGTGTGAGTGTGATCTTCTTCACTTCATCGCATTCGGCGCAGCAGTTCCTATGCCC 908  
Db 1955 TCACCGTCTGGGCCACGCTTATGCGGTGTCAATGGCCATCGGTTCGCAATTTGTCGTA 1896  
QY 909 TTTACTTGGTTGCGCGCAAGCGCAGCATTTGATGCAAGTCAACCGCTGCTCCAGTGCCTG 968  
Db 1895 CGGTGATGCTTCAITTCATCCCGCTAAACGTGACGCGCAGCGGTGCCCC----- 1844  
QY 969 CAGGAACACCAAAACCGCAAGCAGCAGCAGCGCAAGATTTTCAACGATTCACCATCA 1028  
Db 1843 -----GAAGAGAAACACACAGAGTTATTACACCACTGAGCAGGCGGTA 1798  
QY 1029 TCCAGGCACTTTGACCGGTGAAGCTATTGACCTGACGAGCTGACGATGCAATGTTG 1088  
Db 1797 TCTGTTCACCGATGACGAGAGATTTGTGCTCATTCACGCTGCTGATACACGTTG 1738  
QY 1089 CCAGCGGAAAGCTTGGCTCGGCGGTGGCATCGTCCCAACCAAGGCGGAGTTAGTTCTC 1148  
Db 1737 CCAGTGGCTGTTGGTAAAGTATTGCCATTCGCTCGCTCGGTGAGTGCCTCTC 1678  
QY 1149 CGGTGAGTGAAGATTTGGTGGCATTCCTATCTGGCCATGCTTTCGAGTTCGACCA 1208  
Db 1677 CGGTTGCGGTCGAATTTGCTTGTGTTGCGCCACATTACA-----CGCCATTGGCATTG 1624  
QY 1209 AGGCTGAGATGTTCCATGTTGGATATCTTGTGATGATGCTTTCGACAGTAAAC 1268  
Db 1623 AGTCAGATGATGG-----TGTGGAGATCCTGATTCATGTCGGTATCGACACCGTAAC 1570  
QY 1269 TCAACGCGACGCACTTTAAACCGCTCAAGACGAGGCGCATGAAGTCAAGAGGAGGAGC 1328  
Db 1569 TGAACGCAAAATCTTTTCGCTCAGCTCAACGTTGCAAGGTCAATACAGCGATC 1510  
QY 1329 TCGTGTGAATTCGATATGATGCAATTAAGGTGAGCTGAGTATGAGGTAAACAGCCGA 1388  
Db 1509 GCGTATTTCTTTGATATCCCTGCTATTTCGGAGGCGCGGATTTGATCTGACGAGCCGG 1450  
QY 1389 TTGTTGTTTCGAATTAACAGAAAACCGGACCTGTAAACACTTACGTTTGGCGAATG 1448  
Db 1449 TATTATCATGTAATAGCGATGATTTTACGGAAGTATATCCCAACGCGCAGCGCATAA 1390  
QY 1449 AAGCGGAGCAACCTGCTCAACGTC 1474  
Db 1389 GCGCAGTGAACCGCTGTTATCCATC 1364

RESULT 11

US-08-920-827-17/c

; Sequence 17, Application US/08920827  
; Patent No. 5770375  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotsugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,827  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3615 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
; STRAIN: Clinical Isolate EC-24  
; US-08-920-827-17

Query Match 7.2%; Score 109.6; DB 1; Length 3615;  
Best Local Similarity 46.3%; Pred. No. 4.4e-23;  
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;

QY 312 TCACCTCCAGTGTGACGTTGCTGCTCACCGGATTCCTTACATTCATGCGCATGGCCCCAG 371  
Db 2495 TCACACCATTTGTATGTCTGATGTTATCACCCGTACCTTTCTGTGGTGGGCGCG 2436  
QY 372 CAATCGCGTGGTGGCGGATGTGCTGGCACACGCTCTACAGGACCTTATGATTTGGTG 431  
Db 2435 TATCAACCTGGATAAGCAACTGATTCGCGCGGTTATCTCTGGCTTATCAGGCGGTT 2376  
QY 432 GTCCAGTGGCGGCTGTGCTCTTCGGTCTGTCTACTCACCATTGTCATCATGTGCTGC 491  
Db 2375 CTGCATTTGCGGCGCGGTAATGGCGGCTTCTGGCAATCTTCGTCATGTTCCGACTGC 2316  
QY 492 ACCAGTCTCTCCGCCAATTGAGCTGGAGCTGTTAAACAGGTTGATGCTTTCATCTTCG 551  
Db 2315 ACTGGGCGCTGTGCGGCTGTATCAATACTTACCGTGTGGCTACGACACATGA 2256  
QY 552 CAACGGCATCTATG---GCTAATATGCCCCAGGTCGCGCATGTTTGGCAGTGTCTTCC 608  
Db 2255 TCCCGCTGTTAATGCCGCCATTATGGCGAGTCCGGCGCGCTTCGCGCTCTCTCT 2196  
QY 609 TGGCGAAGAGTGAAGAAGCTCAAGGCGCTTCAGGTCCTTCAGGTGTCTCCGCTGTTCTTG 668

2195 GCAGACGCGATGCGCAAAAAAGTGTGGCGGATCAGCGCGGTTGACGAGTCTGTTG 2136  
669 GTATTACGAGCGCTGCGATCTCGGTGTGAACCTTGCCTCGCGTGCCTGCTCTTCTATCG 728  
2135 GTATACCGACGACGCGTATATGGGTCAACCTGCGCGTAAGTACCCCTTTGTTATCG 2076  
729 GTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCACCTTTTAATATCAAGCGCAGTTG 788  
2075 CCTGTATCAGTGGGCTTTGGGGCCACCATTAATTTGGTACGCGCAACGAAAGTCTACT 2016  
789 CGTTGGCGCTGCGAGTTCTTGGGTGTTGTTCTATTGATGCTCCAGATATGCTATGT 848  
2015 CCTTGTGTCAGAGTATTTTCACTTCATGCAAAACCAATCCCGTCAACGGGAATGATT 1956  
849 TCTTGTGTGTGCGATTTTACCTTCTTATCGATTCGATTCGCGCGAGGATTTGCTTATGGCC 908  
1955 TCACCGTCTGGGCGGCTTATTTGGCGGTCTATTCGCCATCGTTGGCGATTTGTGCGTA 1896  
909 TTTACTTGTGTCGCGCAACGCGAGCATGATCCAGATCAACCGCTGCTTCCAGTGCCTG 968  
1895 CGGTGATGCTTATTTTCACTACCGCTAAACGTCAGCCAGCGAGGTGCCCC----- 1844  
969 CAGGACGACCAAGCCGAGAGAGACCCCGCAATTTTCAAGATTCACCATCA 1028  
1843 -----GAAGAGAAACACAGAGTTATTACACCATCGAGCGGCGTA 1798  
1029 TCCAGGACCTTTTACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGCCATCTTTG 1088  
1797 TCTGTTACCGATGACGCGAGAGATTTGTGCTGCTATTACGTCGCTGATACACGCTTTG 1738  
1089 CAGCGGAAAGTTGCTGCGGGTGGCCATCGTCCCAACCAAGGGCGAGTTAGTTTCTC 1148  
1737 CAGTGGCGCTGTTGGGTAAGGTATTGCCATTCGCCCTCGGTTGTTGAGTGCCTTCTC 1678  
1149 CGGTGAGTGAAGATTTGTTGGCATTCCTCATCTGGCCATGTTTCGAGTTTCGACCA 1208  
1677 CGGTTGCGGGTGAATGCTTGTGTTGCGCCACATTAACA-----CGCCATTGGCATTG 1624  
1209 AGGCTGAGATGTTTCAATGTTGATATCTGATGACATGTTTTCGACAGTAAAC 1268  
1623 AGTCAGATGATG-----TGTTGAGATCTGATTCATGTCGTTATCGACACGTAAC 1570  
1269 TCAACGCGACGACTTTAACCCGCTGAAGAAGCAGCGCGATGAAGTCAAGCAGGGAGC 1328  
1569 TGGACGGCAAAATCTTTTCGCTCAGTCAACGTTGTTGCAAGTCAATACAGCGATC 1510  
1329 TGCTGTGTGAATTCGATATTGATGCTTAAAGCTGAGGTTATGAGGTAACACGCGGA 1388  
1509 GGCTGATTTCTTTTGAATPCCCTGCTATTTCGAGGCGCGGATTTGATCTGACAGCGCG 1450  
1389 TTGTTGTTTTCGAATTACAGAAACCGGACCTGTAACACTTACGTTTGGGCGAAATTG 1448  
1449 TATTAAATCAGTAATAGCATGATTTTACGAGCTATTACCCACGCGCAGCGCGCATTA 1390  
1449 AAGCGGAGCCAACTGCTCAACGTC 1474  
1389 GCGAGGTGAACCGCTGTTATCCATC 1364

RESULT 12  
US-08-921-177-17/c  
; Sequence 17, Application US/08921177  
; Patent No. 5798211  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3615 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
STRAIN: Clinical Isolate EC-24  
US-08-921-177-17

Query Match 7.2%; Score 109.6; DB 1; Length 3615;  
Best Local Similarity 46.3%; Pred. No. 4.4e-23;  
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;  
QY 312 TCATCTCCAGTGTGAGTGTGCTCACCGGATTCCTTACATTCATCGCATTCGCCAG 371  
Db 2495 TCACACCATGCTATGCTGATGTTATACACCGTACCTTCTGCTGGTGGGCGCG 2436  
QY 372 CAATCGCTGGTGGCGGATGCTGGCACACGGTCTACAGGAGCTTTATGATTTCCGTTG 431  
Db 2435 TATCAACCTGGATAAGCGAACTGATTGCGCGGTTATCTCTGGCTTTATCAGGCGGTT 2376  
QY 432 GTCCAGTGGCGGCTGCTCTTCGCTGCTGCTTACTCACCATCGTCATCATCTGCTCG 491  
Db 2375 CTGCAATTTGGCGCGCGGTAATGCGCGGCTTCTGCGCAATCTTCGTCATGTTCCGACTG 2316  
QY 492 ACCAGTCTCTCCGCAATTCAGCTGGAGCTGTTTAAACAGGCTGGATCTTTCATCTTCG 551  
Db 2315 ACTGGGCGCTGCTGGCGGTTGTTATCAATACTTACCGTGTGGCTACGACACATGA 2256  
QY 552 CAACCGCATCTATG---GCTAATATCGCCAGGTTGCGGCATGTTTGGCAGTGTCTTCC 608  
Db 2255 TCCCGCTGTTAATGCCCGCATTAATGCGCGAGTTCGCGCGGCTCGCGGCTTCTCT 2196  
QY 609 TGGCGAAGTGAAGAGCTCAAGGCGCTTCAAGGCTTCAAGGCTTCAAGTGTCTTCTTG 668  
Db 2195 CGCAACGCGATGCGCAGAAAGTGGTGGCGGATCAGCGCGGTTGACGAGTCTGTTG 2136  
QY 669 GTATTACGAGCTGCGCATCTTCGCTGTGAACCTTCGCTGCGCTGGCGCTTCTTCATCG 728  
Db 2135 GTATCAGCAACGCGGATATGCGGTCAACCTTCCGCGTAAGTACCCCTTTGTTATCG 2076  
QY 729 GTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCGACTCTTTAATATCAAGGAGTTG 788  
Db 2075 CCTGTATCAGTGGGCTTTGGGGCGCCACCATTAATTTGGCTACGCGCAACGAGTCTACT 2016

QY	789	CGTTGGCGCGCTGCAGGTTTCTTGGGTGTTGTTTCTATTGATGCTCCAGAAATGTCATGT	848
Db	2015	CCITTGGTTTGGCAAGTATTTTTCACCTTCATGCAAAACCATCCGTCACACGGAAATGATTT	1956
QY	849	TCITTGGTGTGTGAGATTGTTTACCTTCTTCATCCATTCCGGCGCAGCGATTGCTATTGGCC	908
Db	1955	TCACCGTCTGGCGCAGCGTTTATTGGCGGTGTCATTGCCATCCGTTCGGCAATTTGTCGTA	1896
QY	909	TTTACTTTGGTTTCGCCGAAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCCAGTGCCTG	968
Db	1895	CGGTGATGCTTCAATTTTCATCACCGCTAAACGTTCAGCCAGCGAGGTGCCCC-----	1844
QY	969	CAGGAACGACCAAGCCGAAGCAGAACGCCAGAAATTTTCAACAGATTCCACATCA	1028
Db	1843	-----GAAGAGAAACACCGAGGTTATTACCAACCTGAGCAGGGCGGTA	1798
QY	1029	TCCAGGCACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTTCAGCGATGCCATGTTTG	1088
Db	1797	TCCTGTTACCGAATCAGCGGAGAGATTGTGTGCTCATTCACGTCGCTGATACCACGTTTG	1738
QY	1089	CCAGCGGAAGCTTGGCTCGGGGTTCGCCATCGTCCCAACCAAGGGCGAGTTAGTTTCTC	1148
Db	1737	CCAGTGGCGCTTTCGGGTAAAGTATTGCCAATTCGCCCCGTGGTTGTGAAGTCGCTTCTC	1678
QY	1149	CGGTGAGTGGAAAGATTGTGTGGTCATTCCCATCTGGCCCATGCTTTGCGAGTTTCGCACCA	1208
Db	1677	CGGTTGCGGGTTCGAATTGCTTCGTTGTTCGCCACATTACA-----CGCCATTGGCATTG	1624
QY	1209	AGGCTGAGGATGGTTTCCAATGTGGATATCTTGATGCAATGTGTTTCGACACAGTAAACC	1268
Db	1623	AGTCAGATGATGG-----TGTCGAGATCCCTGATTTCATGTCGGTATCGACACCGTAAAC	1570
QY	1269	TCAACGCGACGCACTTTTAAACCGGCTGAAGACGAGGCGGATGAAGTCAAAAGCAGGGGAGC	1328
Db	1569	TGGACGGCAAAATTCCTTTTCGCTCAGCGTCAACGTGGGTGAACAAGGTCAATACAGGGGATC	1510
QY	1329	TCCTGTGTGAATTGCATATTGATGCATTAAAGGCTGCAGGTTATGAGGTAAACACGCCGA	1388
Db	1509	GGCTGATTTCTTTTGATATCCCTGCTATTTCGAGAGCGGATTTGATCTCAGCAGCGCGG	1450
QY	1389	TTGTTTGTGTTTCCAAATPACAGAAACCGGACCTGTAAACACTTACGTTTGGGCGAAATG	1448
Db	1449	TATTAAATCAGTAATAGCGATGATTTTACGGACGTATTACCCTCCGCGCACCGCGCAGATAA	1390
QY	1449	AGCGCGAGCCAACTGTCTCAACGTC	1474
Db	1389	GGCGAGGTGAACCGCTGTTATTCATC	1364

## RESULT 13

US-08-362-577C-17/C	QY	669	GTATTACGGAGCGCTCGGATCTTTGGGTGGAACCTTCGCTCGCTGGCGCGTGTCTTCATCG	728
Sequence 17, Application US/08362577C				
Patent No. 5807673				
GENERAL INFORMATION:				
APPLICANT: Ohno, Tsuneya	QY	729	GTATCGGTACCGCAGCATATCGGTGGCGCTTTCGATTGTCACACTCTTTAATATATCAAGCGCAGTTG	788
APPLICANT: Matsuhisa, Akio				
APPLICANT: Uehara, Hirotosugu				
APPLICANT: Eda, Soji	Db	2075	CCTGTATCAGTGGGGCTTTGGGGGGCCACCATTATTTGGCTACGGCGCAAACGAAAGTCTACT	2016
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease				
NUMBER OF SEQUENCES: 25	QY	789	CGTTGGGGCGCTGCAGGTTCTTTGGGTGTTTCTTATATGATCGTCCAGATATGTCATGTT	848
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun	Db	2015	CCTTTGGTTTGGCCAACTATTTTTCACCTTCATGCAAAACCATCCCGTCAACGGGAATGATT	1956
STREET: 6300 Sears Tower, 233 South Wacker Drive				
CITY: Chicago	QY	849	TCCTGTGTGTGTCAGTTGTTACCTTCTTCATGCAATTCGGCGCAGCGAATTCGTTATATGGCC	908
STATE: Illinois				
COUNTRY: United States of America	Db	1955	TCACCGTCTGGGCCAGCGTTATTTGGCGGTGTCATTCGCCATCGTTGGCCATTTGTCGHA	1896
ZIP: 60606-6402				
COMPUTER READABLE FORM:	QY	909	TTTACTTTGGTTTCGCCCAACGGCAGCAATTCAGATCCAGATGCAACCGCTGCTCCAGTGCCTG	968
MEDIUM TYPE: Floppy disk				
COMPUTER: IBM PC compatible	Db	1895	CGGTGATGCTTCATTTTCATCACCGCTAAACGTCAGCCAGCGAGGTGGCCCC	1844
OPERATING SYSTEM: PC-DOS/MS-DOS				
SOFTWARE: PatentIn Release #1.0, Version #1.25	QY	969	CAGGAACGCCAAAGCCGAGCAGACAGACCCCGCAGATTTTTCAAAACGATTTCCACCATCA	1028
CURRENT APPLICATION DATA:				

Db 1843 -----GAGAGAAACACAGAGGTTATTACACACCTGAGCAGGCGGTA 1798  
QY 1029 TCAGGACCTTTGACCGTGAAGCTATTGCACCTGAGCAGCGTCAGCATGCGCATGTTG 1088  
Db 1797 TCTGTTTACCGATGACGGGAGAGATTGTGCGCTCATTTACGTCGCTGATACACAGTTG 1738  
QY 1089 CCAGCGGAAGCTTGGCTCGGCGTGGCCATGCTGCCAACCAAGGGGCGATTAGTTTC 1148  
Db 1737 CCAGTGGCCTGTTGGGTAAAGGTAATGGCCATTCTGCCCTCGTGGTGAAGTGGTCTC 1678  
QY 1149 CGGTGAGTGAAGAAATTGTTGGCATTTCCCATCTGCCCATGCTTTCGAGTTGACCA 1208  
Db 1677 CGGTTGGGGTCAATGCTTCTGTTGTCGCCAATTACA-----CGCATTTGGCATTG 1624  
QY 1209 AGCTGAGGATGTTCCAAATGTGGAATCTTGTGACACATTTGTTGACACAGTAAAC 1268  
Db 1623 AGTCAGATATG-----TGTGGAGATCTGATTCTGCTGCTGATCGATCGACCGTAAAC 1570  
QY 1269 TCAACGGCAGCACTTTAAACCCGCTGAAGAGCAGGGCGATGAGTCAAGAGAGGGAGC 1328  
Db 1569 TGAACGGCAAAATCTTTTCGCTCACGTCACGTCACGTTGGGTGACAAAGTCAATACAGGCGATC 1510  
QY 1329 TGTGTGTGAATTCGATATTGATGCCATTAAAGCTGACGTTATGAGTTAGGTAACACAGCGCA 1388  
Db 1509 GGCTGATTTCTTTGATATCCCTGCTATTTCGAGGCGGATTTGATCTGACACGCCG 1450  
QY 1389 TTGTTGTTTGAATACAGAAACCGGACCTGTAAACACTTTACGGTTTGGGCGAAATTG 1448  
Db 1449 TATTATCATGTAATACGATGATTTTACGAGCTATTACCCACGCGACGCGCGCATAA 1390  
QY 1449 AAGCGGAGCAACTGCTCAAGTTC 1474  
Db 1389 GCGCAGGTGAACCGTGTATTCCATC 1364

## RESULT 14

US-08-920-828-17/c  
; Sequence 17, Application US/08920828  
; Patent No. 585398  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,828  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3615 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
; STRAIN: Clinical isolate EC-24  
; US-08-920-828-17

Query Match 7.2%; Score 109.6; DB 2; Length 3615;  
Best Local Similarity 46.3%; Pred. No. 4.4e-23;  
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;  
QY 312 TCATCTCAGTGTGAGTGTGCTGCTCACCGGATTCCTTACATTCATCGCATTCGGCCAG 371  
Db 2495 TCACACCAATTGCTATCTCTGATGTTTATCACCCCTCACCTTCTGCTGGTGGGCGGC 2436  
QY 372 CAATGGCTGGGTGGCGGATGCTGGGCACAGGTCTACAGGACATTTATGATTTCCGTTG 431  
Db 2435 TATCAACCTTGGATTAAGCGAACTGATTGCGCGCGGTTATCTCTGGCTTTATCAGCGGTT 2376  
QY 432 GTCCAGTCGGCGGTCTGCTCTTCGCTGCTGCTTACTACCAATCGTCATCATCTGCTCTG 491  
Db 2375 CTGCATTTGCGGCGCGGTAAATGGCGCGCTTCTGGCAAAATCTTCTGTCATGTTGCGACTG 2316  
QY 492 ACCAGTCTTCCCGCAATTGAGCTGGAGCTGTTTAAACAGGGTGGATCCTTCATCTTCG 551  
Db 2315 ACTGGGCGCTGTCGCGGTGTGTATCAATAAATTACACCTGCTGGGTACGACACATGA 2256  
QY 552 CAACGGCACTATG--GCTAATATCGCCAGGGTGGCATGTTTGGCAGTCTTCTTC 608  
Db 2255 TCCCGCTGTAAATGCCGCCATTATGGCGAGGTGGGGCGCGCTCGCGCTTCTCTCT 2196  
QY 609 TGGCGAAGAGTGAAGGCTCAAGGCGCTTGACAGTGTCTTCAAGTGTCTCCGCTGTCTTG 668  
Db 2195 GCGAAGCGATGCGCAGAAAAAAGTGTGGCGGATCAGCGCGTGTGACAGTCTGTGTTG 2136  
QY 669 GTATTACGAGCTCGCATCTTTCGGTGTGAACCTTCGCTGGCTGGCGCTTCTTCATCG 728  
Db 2135 GTATCACCGAACCGAGCGGTATATGGCTCACTGCGCGGTAAAGTACCCCTTGTATCG 2076  
QY 729 GTATCGGTACCGCAGCTATCGGTGGCGCTTGAATTGCACTCTTTAATATCAAGGAGTTG 788  
Db 2075 CCTGTATCAGTGGGCGCTTGGGGGCCACCAATTATTTGGCTACGCGCAAAACGAAAGTCTACT 2016  
QY 789 CGTTGGGCGCTGCAAGTCTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATATGGTCAATGT 848  
Db 2015 CTTTGGTTTGCAGATATTTTACCTTCATGCAAAACCATCCCGTCAACGGGAATTGATT 1956  
QY 849 TCTTGGTGTGCAAGTGTATTACCTTCTTCATCGCATTCGGCGCAGCATTCCTTATGGCC 908  
Db 1955 TCACCGCTGCGGCCAGCGTTATTGCGGTGTCATTGCCATCGGTTCGCAATTTGTCGTA 1896  
QY 909 TTTACTTGTTCGCGCAGCAGCGAGCATGATGATGATGCAACCGCTGCTCCAGTGGCTG 968  
Db 1895 CGGTGATGCTTTCATTTTATCATCCGCTAAAGCTACGCGCAGCGCGGTTGCCCTC----- 1844  
QY 969 CAGGAACGACCAAGCCGAGCAGAACCCCGCAGAAATTTTCAACGATTTCCACCATCA 1028  
Db 1843 -----GAGAGAAACACAGAGGTTATTACACCTGAGCAGGCGGTA 1798  
QY 1029 TCAGGACCTTTGACCGGTGAAGCTATTGACATGAGCAGCGTCAGCGCATGCCATGTTG 1088  
Db 1797 TCTGTTTACCGATGACCGGAGAGATTGTGCTGCTCATTCAGTGTGCTGATACACAGTTG 1738  
QY 1089 CCAGCGGAAGCTTGGCTCGGCGGTTGCCATCGTCCCAACCAAGGGGCGATTAGTTTCTC 1148  
Db 1737 CCAGTGGCCTGTTGGGTAAAGGTAATGGCCATTCTGCCCTCGTGGTGTGAGTGGTCTC 1678  
QY 1149 CGGTGAGTGAAGATTTGTTGGCTATTCATTTCCCATTTGCGCATTTCTTCGAGTTCGACCA 1208



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 15:50:57 ; Search time 744 Seconds  
(without alignments)  
10523.989 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 cctatggatctgcgcgtt.....gttgaaacctgagtgttcg 1527

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	1527	100.0	1527	16	US-10-450-055-9
2	1515	99.2	3309400	9	US-09-738-626-1
3	1492	97.7	1983	9	US-09-738-626-2904
4	177.8	11.6	975	9	US-09-974-300-829
5	151.2	9.9	5840	9	US-09-070-927A-31
6	137.2	9.0	1380	9	US-09-974-300-857
7	135.4	8.9	3037	16	US-10-398-221-3599
8	134.6	8.8	684707	16	US-10-398-221-9
9	134.6	8.8	3011208	16	US-10-398-221-2058
10	126.6	8.3	2127	17	US-10-474-776-160
11	126.6	8.3	3895	8	US-08-961-527-201
12	126.6	8.3	3895	16	US-10-158-844-201

13	113.4	7.4	1884	9	US-09-815-242-9369	Sequence 9369, App
14	107.8	7.1	1185	17	US-10-474-776-152	Sequence 152, App
15	107.8	7.1	9769	8	US-08-961-527-30	Sequence 30, Appl
16	107.8	7.1	9769	16	US-10-158-844-30	Sequence 822, App
17	105.6	6.9	1014	9	US-09-974-300-822	Sequence 822, App
18	99	6.5	1098	9	US-09-974-300-734	Sequence 734, App
19	98	6.4	489	9	US-09-974-300-857	Sequence 857, App
20	93.6	6.1	1854	16	US-10-398-221-712	Sequence 712, App
21	93.6	6.1	1854	16	US-10-398-221-2722	Sequence 2722, App
22	93.6	6.1	1854	16	US-10-282-122A-24691	Sequence 24691, A
23	87.8	5.7	1137	16	US-10-282-122A-34025	Sequence 34025, A
24	86.2	5.6	30246	8	US-08-781-986A-56	Sequence 56, Appl
25	86.2	5.6	30246	16	US-10-329-624-56	Sequence 56, Appl
26	81.6	5.3	1854	16	US-10-398-221-1040	Sequence 1040, App
27	81.6	5.3	1854	16	US-10-398-221-2857	Sequence 2857, App
28	78.8	5.2	12278	16	US-10-398-221-3886	Sequence 3886, App
29	77.4	5.1	444	9	US-09-974-300-877	Sequence 877, App
30	76.8	5.0	1368	9	US-09-974-300-893	Sequence 893, App
31	76.4	5.0	465	8	US-08-781-986A-1317	Sequence 1317, App
32	76.4	5.0	465	16	US-10-329-624-1317	Sequence 1317, App
33	74.2	4.9	474	9	US-09-070-927A-639	Sequence 639, App
34	73.2	4.8	1357	16	US-10-282-122A-21444	Sequence 21444, A
35	72.4	4.7	1953	16	US-10-282-122A-38846	Sequence 38846, A
36	72.4	4.7	1956	16	US-10-282-122A-36728	Sequence 36728, A
37	70.4	4.6	9797	9	US-09-070-927A-550	Sequence 550, App
38	69.8	4.6	1353	9	US-09-974-300-799	Sequence 799, App
39	69.4	4.5	429	9	US-09-974-300-5182	Sequence 5182, App
40	69.2	4.5	1953	16	US-10-282-122A-39556	Sequence 39556, A
41	68.2	4.5	348	9	US-09-974-300-5313	Sequence 5313, App
42	67.2	4.4	2028	16	US-10-282-122A-16373	Sequence 16373, A
43	67	4.4	2181	17	US-10-474-776-65	Sequence 65, Appl
44	67	4.4	8494	8	US-08-961-527-163	Sequence 163, App
45	67	4.4	8494	16	US-10-158-844-163	Sequence 163, App

#### ALIGNMENTS

RESULT 1  
US-10-450-055-9  
; Sequence 9, Application US/10450055  
; Publication No. US20040043953A1  
; GENERAL INFORMATION:  
; APPLICANT: BASF Aktiengesellschaft  
; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium  
; FILE REFERENCE: 936 2000  
; CURRENT APPLICATION NUMBER: US/10/450,055  
; CURRENT FILING DATE: 2003-06-10  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Vers. 2.0  
; SEQ ID NO 9  
; LENGTH: 1527  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (101)..(1504)  
; OTHER INFORMATION: RXS00315  
US-10-450-055-9

Query Match	100.0%	Score 1527;	DB 16;	Length 1527;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1527;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CTCATGGCATCTGCGCGCTTCGCGTTCCTGCGAGTGTGTGTTTCCACCGCAACCAAG	60	
Db	1	CTCATGGCATCTGCGCGCTTCGCGTTCCTGCGAGTGTGTGTTTCCACCGCAACCAAG	60	
Qy	61	CGTTTCGGCGGCATAGTTCCTGGGCGCCCGCTATTGATGGCGATGTTGTTCCCGAG	120	
Db	61	CGTTTCGGCGGCATAGTTCCTGGGCGCCCGCTATTGATGGCGATGTTGTTCCCGAG	120	
Qy	121	CTTGGTGACGGCTACGACGTGGCGGCCACCATGGCTCGGCGGAATGCCAATGTGGTC	180	

Db 121 CTTGGTGAACGGCTACGAGTGGCGCCACCATGGCTGCGGCGAATGCAATGTGTC 180  
QY 181 CTTGTTTGGTTAGATTGCTCCCAAGCCGTTACAGGGCAACCGTGTCTCTGTGCTGGT 240  
Db 181 CTTGTTTGGTTAGATTGCTCCCAAGCCGTTACAGGGCAACCGTGTCTCTGTGCTGGT 240  
QY 241 GGTTCCTTGTGATTCTGCAACGATCGAAGATTCTCTGCAACAGGACTCAAGGCACTGC 300  
Db 241 GGTTCCTTGTGATTCTGCAACGATCGAAGATTCTCTGCAACAGGACTCAAGGCACTGC 300  
QY 301 AGACTTCCTGATCACTCCAGTGTGAGTGTCTCTCTCCTCAGCGGATTCCTTACATTCATCGC 360  
Db 301 AGACTTCCTGATCACTCCAGTGTGAGTGTCTCTCTCCTCAGCGGATTCCTTACATTCATCGC 360  
QY 361 CATTGGCCCAAGCAATCGCTGGGTGGCGATGTCTGGCACACGCTCTACAGGACATTTA 420  
Db 361 CATTGGCCCAAGCAATCGCTGGGTGGCGATGTCTGGCACACGCTCTACAGGACATTTA 420  
QY 421 TGATTTCGTTGGTCCAGTCCGCGGTCTGCTCTTCGCTCTGCTCTACTCACCAATCGTCAT 480  
Db 421 TGATTTCGTTGGTCCAGTCCGCGGTCTGCTCTTCGCTCTGCTCTACTCACCAATCGTCAT 480  
QY 481 CACTGTGCTGCACAGTCTCTCCGCGCAATTTGAGCTGGAGCTGTTTAAACCAAGGCTGGATC 540  
Db 481 CACTGTGCTGCACAGTCTCTCCGCGCAATTTGAGCTGGAGCTGTTTAAACCAAGGCTGGATC 540  
QY 541 CTTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGCTGGCGATGTTTGGCAAT 600  
Db 541 CTTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGCTGGCGATGTTTGGCAAT 600  
QY 601 GTTCTTCTGCGAAGAGTGAAGCTCAAGGCTCTGAGGCTCTGAGTGTCTCAGTGTCTCGC 660  
Db 601 GTTCTTCTGCGAAGAGTGAAGCTCAAGGCTCTGAGGCTCTGAGTGTCTCAGTGTCTCGC 660  
QY 661 TGTTCCTGATTAACGAGCTCTGATCTCTCGGTGTGAACCTTCGCTCTGCGTGGCGGTT 720  
Db 661 TGTTCCTGATTAACGAGCTCTGATCTCTCGGTGTGAACCTTCGCTCTGCGTGGCGGTT 720  
QY 721 CTTTCATCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATGACATCTTTAATATCAA 780  
Db 721 CTTTCATCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATGACATCTTTAATATCAA 780  
QY 781 GGCAGTTGGTGGCGCTGCAAGTCTTCTGGTGTGTTCTTATGATGCTCCAGATAT 840  
Db 781 GGCAGTTGGTGGCGCTGCAAGTCTTCTGGTGTGTTCTTATGATGCTCCAGATAT 840  
QY 841 GGTGATGTTCTGTGTGTGCAAGTGTGTTACCTTCTTCATCGCATTCGCGCAGCGATTCG 900  
Db 841 GGTGATGTTCTGTGTGTGCAAGTGTGTTACCTTCTTCATCGCATTCGCGCAGCGATTCG 900  
QY 901 TTATGCGCTTTACTTGGTTCGCGCAACGCGAGCATTTGATCCAGATGCAACCGCTGCTCC 960  
Db 901 TTATGCGCTTTACTTGGTTCGCGCAACGCGAGCATTTGATCCAGATGCAACCGCTGCTCC 960  
QY 961 AGTGCTGCGAAGACGACCAAGCCGAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1020  
Db 961 AGTGCTGCGAAGACGACCAAGCCGAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1020  
QY 1021 CACCATATCCAGGACCTTTGACCGGTGAAGCTATTCACCTGAGCAGCTGAGCGATGC 1080  
Db 1021 CACCATATCCAGGACCTTTGACCGGTGAAGCTATTCACCTGAGCAGCTGAGCGATGC 1080  
QY 1081 CATGTTTGCAGGGAAGCTGTGCTGGGGGTGTCATCGTCCCAACCAAGGGGAGTT 1140  
Db 1081 CATGTTTGCAGGGAAGCTGTGCTGGGGGTGTCATCGTCCCAACCAAGGGGAGTT 1140  
QY 1141 AGTTTCTCGGTGAGTGAAGATCTGTGTGGCAATTCCTGCTGCGCATCTTTCGCAAT 1200  
Db 1141 AGTTTCTCGGTGAGTGAAGATCTGTGTGGCAATTCCTGCTGCGCATCTTTCGCAAT 1200  
QY 1201 TCGCACCAAGGCTGAGGATGGTTCCAAATGTGGATATCTTGTATGACATGTTGGTTTCACAC 1260

Db 1201 TCGCACCAAGGCTGAGGATGGTTCCAAATGTGGATATCTTGTATGACATGTTGGTTTCGACAC 1260  
QY 1261 AGTAAACCTCAAGCGCAGCAGCTTTTAAACCCGCTGAAGAACAGCAGGCGATGAAGTCAAGC 1320  
Db 1261 AGTAAACCTCAAGCGCAGCAGCTTTTAAACCCGCTGAAGAACAGCAGGCGATGAAGTCAAGC 1320  
QY 1321 AGGGAGCTCTCTGTGTGAATTCGATATTTGATGTCATTAAGGCTGAGGTTATGAGGTAAC 1380  
Db 1321 AGGGAGCTCTCTGTGTGAATTCGATATTTGATGTCATTAAGGCTGAGGTTATGAGGTAAC 1380  
QY 1381 CAGCCGATTTGTTTTCGAATTTACAGAAAACCGGACCTGTAAACACTTACGTTTGGG 1440  
Db 1381 CAGCCGATTTGTTTTCGAATTTACAGAAAACCGGACCTGTAAACACTTACGTTTGGG 1440  
QY 1441 CGAAATTTGAAGCGGAGCCCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTGCCAGCAAC 1500  
Db 1441 CGAAATTTGAAGCGGAGCCCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTGCCAGCAAC 1500  
QY 1501 ACCATAAGTTGAAACCTTTGAGTGTTCG 1527  
Db 1501 ACCATAAGTTGAAACCTTTGAGTGTTCG 1527

## RESULT 2

US-09-738-626-1/c

; Sequence 1, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAOKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; PRIORITY FILING DATE: 2000-12-18

; PRIORITY APPLICATION NUMBER: JP 99/377484

; PRIORITY FILING DATE: 1999-12-16

; PRIORITY APPLICATION NUMBER: JP 00/159162

; PRIORITY FILING DATE: 2000-04-07

; PRIORITY APPLICATION NUMBER: JP 00/280988

; PRIORITY FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 1

; LENGTH: 3309400

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-1

Query Match 99.2%; Score 1515; DB 9; Length 3309400;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1526; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CTTATGGCATCTGCGCGCTTCGCGTTCTTGCAGTGTGTTTCCCGCAACCAAG 60  
Db 2811326 CTTATGGCATCTGCGCGCTTCGCGTTCTTGCAGTGTGTTTCCCGCAACCAAG 2811267  
QY 61 CTTTTCGCGGCAATGAGTTCTTGGCGCGCTTGGTATGGCATGTTGTTCCCGAG 120  
Db 2811266 CTTTTCGCGGCAATGAGTTCTTGGCGCGG-CTTATGTAAGCGATGTTCCCGAG 2811208  
QY 121 CTTTGTGAACGGCTACGACGTGGCGCCACCATGCTCGGCGCAATGCCAATGTGTC 180  
Db 2811207 CTTTGTGAACGGCTACGACGTGGCGCCACCATGCTCGGCGCAATGCCAATGTGTC 2811148  
QY 181 CTTGTTGTTTGTAGATGTTGCGCCAGCCGTTACCAAGGCGCCTTCTCTGCTGCT 240





Db 466 GAACAATTCGCCGTAAGTGGCGGAGTCTTATACGAGGGCTTTACGCCGCTCTCGTCATT 525  
QY 482 ACTGGTCTGCACAGCTCTTCCGCCAAATGAGCTGGAGCTGTTT-----AACCAAGGT 535  
Db 526 ACCGGAATGATACATACATCTCTGCGGTGACCTTCAGCTGATCGCTCAAGCTCGGC 585  
QY 536 GATATCTTCATCTCCGACGGAATCTATGGCTPAATATCGCCAGGCTGGCGCATGTTG 595  
Db 586 GGAACATATTTTATGGCGGATGCTCGCGCTGTCATATATGACAAAGGTTTCAGCGCGCT 645  
QY 596 GCAGTCTCTCTCGCGGAGAGTGAAGAGCTCAAGGCGCTTCAGGTGCTTCAGGTGTC 655  
Db 646 GCATGATGTTTATGTCAAAGATGAGAGCGAAGAGGCTGTCCTGACATCCGGAAT 705  
QY 656 TCCTGCTGTTCTGGTATTACGGAGCCTCGCATCTTCGGGTGTGAACCTTCGCCCTGCGCTG 715  
Db 706 TCAGCTTATCTCGGAATTACCGACCGCATGTTTCGGAGTGAATCTGATACAGATTT 765  
QY 716 CCGTCTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCACCTCTTTAAT 775  
Db 766 CCGTTCGTATCGCATGATCAGTTCCGGGATTCGGGTATGTTTCTATTTCTCGCAAGG 825  
QY 776 ATCAAGCAGTTCGTTGGCGCTGCGAGTTCCTTGGGTGTTGTTCTATT 826  
Db 826 GTTTGGCAAGCTCTGTCGGCGTTCGGCGGTACCTGGGAATTTCTCTATT 876

## RESULT 5

US-09-070-927A-31  
; Sequence 31, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillon  
; Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 982  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070,927A  
; FILING DATE: 04-May-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/046,655  
; FILING DATE: 1997-05-16  
; APPLICATION NUMBER: 60/044,031  
; FILING DATE: 1997-05-06  
; APPLICATION NUMBER: 60/066,009  
; FILING DATE: 1997-11-14  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kenley K. Hoover  
; REGISTRATION NUMBER: 40,302  
; REFERENCE/DOCKET NUMBER: PB369  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5840 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-070-927A-31

Query Match 9.9%; Score 151.2; DB 9; Length 5840;  
Best Local Similarity 48.5%; Pred. No. 8.5e-36;  
Matches 525; Conservative 0; Mismatches 533; Indels 24; Gaps 3;  
QY 329 TTGCTGCTCACCGGATTCCTTACATTCATCGCCATTTGGCCCGCAGCAATCGCTGGTGGGC 388  
Db 1850 TTCTTTATTTATGGCACTGTTTACCTTCTTAGCTATTGGTCTTATCGGGAAGGTCATCGGC 1909  
QY 389 GATGCTCTGGCACACGGTCTACAGGACCTTATGATTTTCGGTGTCTCAGTCGCGGCTG 448  
Db 1910 GATTTGTTGGACAGGATACACGCAATTTACGCTTTTAGTCCAAATTAACGAGGTTA 1969  
QY 449 CTCCTCGCTCTGGTCTACTCACCAATTCGTCATCGTCTGCACCGCTCTTCCCGCCA 508  
Db 1970 TTAATGGGTTCTGTTGGCAAGTCTTGGTAAATGTTTGGTATGCAATTTGGGCTTTGTACCA 2029  
QY 509 ATTGAGCTGAGCTGTTTAAACAGGCTGATCCTTCATCTTCGACGCGCACTATGGCT 568  
Db 2030 ATTATGATGTTTAACTTAAACAAGGTGGCATACGATGGTACCGATGTTATTACAGCC 2089  
QY 569 AATATCGCCAGGCTCGGCATGTTTGGCAGTGTCTTCTCGCGCAAGAGTGAAGAGCTC 628  
Db 2090 GTTATGTCACAAGCGCGGCTGCTTTAGCTGCTTTTCTTAACAAAATATGAAACTA 2149  
QY 629 AAGGCTTTGCAAGGTCTTCAAGTGTCTCCGCTGTTCTTGGTATTACGAGCTCGCATC 688  
Db 2150 AAAAGTTTGGCTTTGCTTCAAGTATTACGACTATTTTGGAAATTAATGAACTGTA 2209  
QY 689 TTCGGTGTGAACCTTCGCTCGCTGGCTGCGCTTCTTCATCGGTATCGGTACCGCAGCTATC 748  
Db 2210 TATGGCTGACTTTTACCATTGAAAAAACCATTATTTCGAGCTGTGATTGGTGGCGGTATC 2269  
QY 749 GGTGGCGCTTTGATTCGACTCTTTTAATATCAAGGAGTTCGCTTGGCGCTCGCAGGTTTC 808  
Db 2270 GGTGGTGCATTTGTGCTATGATCAAGTGAAGAACTTTACGTTTGGCTTGGTGTATG 2329  
QY 809 TTGGTGTGTTTCTTATGATGCTCCAGATGTTGTCATGTTCTTGGTGTGTCAGTGT 868  
Db 2330 TTGAGCTTGGCTGGCTTTTATCTCGCAGAGACAAAAGATCTGCACCGATGATTACTGCT 2389  
QY 869 ACCTTCTTTCATCGCATTCGGCGCAGGATTCGCTTATGCTTATGCTTACTTGGTTCGCCCAAC 928  
Db 2390 GCAATGGTGGCGGAATTCGCTTTATCATGCTGTTTGTCTTAAACGTTTGTCTTACGTT-- 2447  
QY 929 GGCAGCATTTGATCCAGATCGCAACCGCTGCTCCAGTGCCTGCGAGGAACGACCAAGCCGAA 988  
Db 2448 -----TTGAAGATCAACCTAATCCAGAACAGCAACTGAAAAACAGAACTGATAAG 2500  
QY 989 CGAGAAGCACCGCGAGAAATTTTCAACGATTCACATCCATCATCCAGGCA---CCTTTGACC 1045  
Db 2501 ATGTTGGCACCTGTAAAAACGAAATCAAGAGACAAAAATTTTATAGCAAGTCCACTTCAA 2560  
QY 1046 GGTGAAGCTATTGCACTGAGCAGGCTCAGCGATGCAATGTTTGCAGCGGAAAGTTGGC 1105  
Db 2561 GGTGAATTTTACCGCTAGAAAAAGTACAACCTGTTTTCAGTTCAGGTGCTTTAGGA 2620  
QY 1106 TCGGCGTTCATGCTGCTCCCAACCAAGGGGAGTTAGTTTCTTCGGTTCAGTGGAAAGATT 1165  
Db 2621 AAAGGTGTTGCAATTTAGCCGACTGAAGCAAACTGATATGACCCGCGAGATGGTAAATC 2680  
QY 1166 GTGGTGGCAATTCCTCATCTGGCCATGCTTTTCAGTTCGACCAAGGCTGAGATGTTTCC 1225  
Db 2681 ACCACATTTATTCGACAGGACATGCTGTTGGCTTGAACCAACAGAGGGC----- 2731  
QY 1226 AATGTGGATATCTTGATGCAATTTGGTTTCGACACAGTAACCTCAACGGCAGCAGCTTT 1285  
Db 2732 ---GTTGAATTTAATGTCATATTTGGCATGATACGGTTCGAATTAGATGTTAAAGGCTTT 2788  
QY 1286 AACCCGCTGAAGAAGCAGGCGGATGAAGTCAAGCAGGAGGAGCTGCTGTGTAATTCAT 1345



QY 853 GGTGTGTCAGTTGTTTACCTTCTTCATCGCATTCGGCGCAGCATGCTTATGGCCCTTA 912  
 Db 1218 CGAATTCGTGGGTAGTTATCGCAATCGTTATCAGCTTCACTCTTGGTTTCATCTTAAC 1277  
 QY 913 CTTGGTTCGCGGCAACGGCAGCAATTGATCAGATGCAACCGCTGCTCCATGCTGCGAGG 972  
 Db 1278 TTATGTA-----GTAGGTTTCAAGATCCAGCTGATGATAGTAGTAGAACA 1322  
 QY 973 AACGACAAAGCCGAAGCAGACGACCGCAGAAATTTTCAAGATTTCCACCATCATCCA 1032  
 Db 1323 ATCTAACACAGTAGAGGCGAAGC-----TTTAAATCGAAGTGAACAATTC 1370  
 QY 1033 GGCACCTTTGACCGGTGAAGCTATTCACATGACGAGCGTCAGCGATGCCATGTTTGCAG 1092  
 Db 1371 AGCTCCAGTAGTTGGCGAAATCGTAACTTTTAGCAGAGCTAAAGACGAACATTTTCATC 1430  
 QY 1093 CGGAAAGCTTGGCTCGGCGGTTCGCAATCGTCCCAACCAAGGGCAGTTAGTTCTCGGT 1152  
 Db 1431 CGGTGCACCTTGGAAAGGTGTTGCAATCATCCCAACTGTTGGACGTGTAGTGTCCAGC 1490  
 QY 1153 GAGTGAAGATTTGGTGGCATTTCCATCTGGCCATGCTTTTCGAGTTTCGACACAGGC 1212  
 Db 1491 AGCAGGAACAGTAACACTATCTTCCCACTGGTCAATCGGTATTAACAATAAG- 1549  
 QY 1213 TGAGGATGGTCCAAATGTGGATATCTTGATGCAATTTGTTTCGACACAGTAACCTCAA 1272  
 Db 1550 -----ACGGTGCAGAAAGTTTAAATCCACATCGGTATGACACCGCTTCAACTGA 1598  
 QY 1273 CGGCGACGACTTAAACCCCTGAGAGCAGGCGGATGAGTCAAGAGCGGGAGCTGCT 1332  
 Db 1599 AGGTAAATTTCTCACAGCACACGTGAACCAAGCGACGTAATCGAAAAAGGTCAATTA 1658  
 QY 1333 GTGTGAATTCGATATTGATGCCATTAAAGCTGCGAGTTATGAGTAAACACCGGATTCT 1392  
 Db 1659 AACTGAATTTGATTTGAAGGCACTAAGAGCTGCTGGATGATGAGTTTACCACTCCAGTAGT 1718  
 QY 1393 TGTTCGAATTAAGAAACCGGACCTGTAAACACT 1429  
 Db 1719 AGTAAACAACTCCAATCAATATTAGACGTATGATT 1755

RESULT 8  
 US-10-398-221-9/c  
 ; Sequence 9, Application US/10398221  
 ; Publication No. US2004018514A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KUNST, Frederik  
 ; APPLICANT: GLASER, Philippe  
 ; TITLE OF INVENTION: Listeria innocua, genome and applications  
 ; FILE REFERENCE: 344 702 - US  
 ; CURRENT APPLICATION NUMBER: US/10/398,221  
 ; PRIORITY FILING DATE: 2003-03-27  
 ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
 ; PRIOR FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: FR 00/12 697  
 ; PRIOR FILING DATE: 2000-10-04  
 ; NUMBER OF SEQ ID NOS: 4025  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 9  
 ; LENGTH: 684707  
 ; TYPE: DNA  
 ; ORGANISM: Listeria innocua  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(end)  
 ; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u

Query Match 8.8%; Score 134.6; DB 16; Length 684707;  
 Best Local Similarity 48.2%; Pred No. 1.3e-29;  
 Matches 537; Conservative 0; Mismatches 534; Indels 42; Gaps 4;

QY 348 TTACATTCATCGCCATTTGGCCAGCAATCGCTGGGTGGCGATGTCTGGCACACCGCTC 407  
 Db 240932 TTACTTTCCTTGTAAATCGGACCAATCGCAACTTGGGCTGTGTCAATTACTTTGGAGCAGGTA 240873  
 QY 408 TACAGGACCTTTATGATTTTCGGTGGTCCAGTCCGCGGTCTGTCTTCTCGTCTGTCTACT 467  
 Db 240872 CAATTTGGGTTTATAACTTAAGCCCGAGTTGTCTGTGTTTAACTTTAGGTGTTTCTGGC 240813  
 QY 468 CACCAATCGTCATCAGTCTGCTGACAGTCCCTTCCGCCCAATTTGAGCTGGAGCTGTTTA 527  
 Db 240812 AAGTATTCGTTATCTTTGGACCTTCACTGGGCTCTTATTCAGTGGCAATCAACAACCTAA 240753  
 QY 528 ACCAGGCTGATCCTTCA---TCTTCGCAACGGCATCTATGCTATATATCGCCAGGCTG 584  
 Db 240752 CAACGCTTGAATCAGATCCAGTACTTGGCAATGATGTTCGGTGTCTTCTTTGGCCAAATCG 240693  
 QY 585 CGGCATGTTTGGCAGTGTCTTCTTCGCGCAAGAGTGAAGAGCTCAAGGGGCTTTGCAGGTG 644  
 Db 240692 GTGCAGTATTAGCAGTATCTTCCAAACTAGAAACAAATAAATCAATCTCTTAGTATTC 240633  
 QY 645 CTTTCAGGTCTCTCGCTGTCTTCTTGGTATTTACGGAGCTCGGATCTTCGGTGTGAACCTTC 704  
 Db 240632 CAGCATTTATTTCCGGTATCTTTTGGTGTAACTGAGCCAGCGATTTACGGGGTTACTTTAC 240573  
 QY 705 GCTGCGCTGGCGGTCTTTCATCGGTATCGGTATCGGTACCGCAGCTATCGTGGCGCTTTCATTG 764  
 Db 240572 CACTGAAAAAACAATTCATCATGAGCTGTATCGAGGTCTGTGGTGGAGGAATCAATTG 240513  
 QY 765 CACTCTTTAATATCAAGGAGTGTGGCTTGGGGGCTGACAGTCTTCTGGGTGTGTTTCTA 824  
 Db 240512 GTTTGTAAAGTCAAAAGTTTATATCATGSGTGGACTTCGAATCTTCGGTCTTCTTAACT 240453  
 QY 825 TTGATGCTCCAGATATGTCATGTTCTTGGTGTGTCAGTGTGTTTACTCTTCTTCATGCAAT 884  
 Db 240452 TCTTCCAAACAGGTCTGTGTTTACAA-----GTGCATTTCTGGTGGTAGTTATCGCAA 240399  
 QY 885 TCGGGCGCAGCGATGCTTATGCGCTTTTACTTGTTCGCGCAACGCGCAGCATTTGATCCAG 944  
 Db 240398 TCGTTATCAGCTTCACTCTTGGTTTCACTTAACTTATGATAGCAGGATTCAGAGATCCAG 240339  
 QY 945 ATGCAACCGCTGCTCCAGTGTCTGCGGAAACGCAACCAAGCCGAAGCAGAACGCCGCGAG 1004  
 Db 240338 CTGAAGCAGTTGTAGAGAAACTAAGTAA-----CAGAGGGCG 240300  
 QY 1005 AATTTTCAACGATTCACCATTCACGACATTCAGGCACTTTGACCGTGAAGTATTTGCACTGA 1064  
 Db 240299 AAACTTTAATCGAACGTGAGACAATTCAGCTCCAGTATTCGCGAAATCGTAACTTTAG 240240  
 QY 1065 GCAGCGTCAGCGATGCCATGTTTTCGCGAGGAAAGCTTGGCTCGGGCGTTGCCATCGTCC 1124  
 Db 240239 CAGAGTAAAGACGAGCAATTTTCATCGGTGCACTTGGAAAAGGTGTAGCAATTTATCC 240180  
 QY 1125 CAACCAAGGGCAGTTAGTTTCTCGGTGAGTGAAGATTTGTGGTGGCATTTCCCATCTG 1184  
 Db 240179 TTACTGTGCGACGCTGTGTAGTCTCCAGCAGCAGGAACAGTAACTATCTTCCCACTG 240120  
 QY 1185 GCCATGCTTTCCAGTTTCGACCAAGCTGAGGATGTTTCCAAATGTTGATATCTTTGATGC 1244  
 Db 240119 GTCATGCAATCGGTATTACAACGAATGA-----CGGTGCAAGATTTAATATCC 240072  
 QY 1245 ACATTTGGTTTCGACACAGTAAACCTCAACGCGCAGCTTTTAAACCGGTGAGAGACGAG 1304  
 Db 240071 ACATCGGTATGATACAGTTCACTAGAGTAAAGTTCTTTCACAGCACTGTTTAAACAG 240012  
 QY 1305 CGGATGAAGTCAAAAGCAGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAGGCTG 1364  
 Db 240011 CGGATGAATCGAAAAAGGTCAATTACTTAAGTGAATTTGATTTGAAGGCATTAAAGCAG 239952  
 QY 1365 CAGGTATGAGGTAAACACGCGGATGTTGTTTTCGAATTAAGAAAAACCGGACCTGTAA 1424  
 Db 239951 CTGGATATGATGTTTACAACTCCAGTAGTAGTAAACAACTCCAAATCAATATTATAGACGTAA 239892  
 QY 1425 ACATTACGGTTTGGCGGAAATTTGAACGGGAG 1457



QY 427 CGGTGTCAGTCCGGGTCGTCTTTCGGTCTGCTACTCACCATCGTCATCTGG 486  
Db |||||  
QY 1077 AGTTAAATGGCTCTCGGTCAAATTTTGGCGCCTCTACGCTCCATTTGTCTACAGG 1136  
Db |||||  
QY 487 TCTGACCACTGCTTCCCGCAATGAGCTGGAGCTGTTT---AACAGGGTGGATCCTT 543  
Db |||||  
QY 1137 TCTGACCACTGATGACCAATGCCATGATACACAAATGATTCGCGATGCTGGTGGCACTGC 1196  
QY 544 CATCTTCGCAACGGCATCTATGCTAATATCGCCACGGGTGGCGCATGTTTGGCAGTGT 603  
Db |||||  
QY 1197 CCTATGCCBAGATGCTCTTCTAATATGCTCAAGGTCAGCGGTGTTTGCCTATTA 1256  
QY 604 CTTCTGCGCAAGAGTGAAGTCAAGGCC---TTGCAAGGTGCTTTCAGGTGCTCCGC 660  
Db |||||  
QY 1257 TTTTCATGCTATCGCATGATGAGCGTGAAGCTCAGGTTTCACTTCTGCAACCAATTCAGC 1316  
QY 661 TGTCTTCTGATTAACGAGCTCGATCTTCGCTGTAACCTTCGCTGCGCTGGCGGCTT 720  
Db |||||  
QY 1317 CTATCTCGGTGTTACAGAACCACTCTTTTGGGGTTAAGCTAAATATATTTATCAAT 1376  
QY 721 CTTTCATCGGTATCGGTACCGCATCTATCGGTGGCGCTTTGATTCGACTCTTTAATATCAA 780  
Db |||||  
QY 1377 TGTTCGTGGATGACTGTTTCAGCCCTTCAGGCAATGTTATCCGTTACTTTTAAATGAAC 1436  
QY 781 GGCAGTGTGGTGGCGCTGCGAGTTCCTTGGGTGTTTGGTGTGTTTCTTATGATGCTCCAGATAT 840  
Db |||||  
QY 1437 TGGCGCTTCTATTTGATCGGTGTTTGGCGGTTTTCAGGATATCTCTTATTCACCTCAATACAT 1496  
QY 841 GTCATGTTCTTGGTGTGTCAGTGTCTTACCTTCCTTCATCGATTCGGCGCAGCGATTCG 900  
Db |||||  
QY 1497 GCTGCCATTTGCGGAACATGCTAGTTGCGATGTTGTTTCCAATGCTC-----TTGAC 1550  
QY 901 TTATGGCTTTTACTTGTTCGCGCAACCGCATCTGATTCAGATGCAACCGCTCTCTCC 960  
Db |||||  
QY 1551 TTTCTTCTCCGAGGCTGCTCTTTTACAAACAGAGGGGATACGAATTCGAGGC 1610  
QY 961 AGTGCCTGAGGAACGCAACAAAGCGAAGCAGACCCCGCAGAAATTTTCAAAAGCATTC 1020  
Db |||||  
QY 1611 AGAATTCGTGTCTCAAGAAAGACAGAAATTTTGAACCATGAACAGTAGAATTTACTTTC 1670  
QY 1021 CACCATCTCCAGGCACCTTTCAGCGGTGAAGCTATTGCACTGACGAGCGTCAGCGATGC 1080  
Db |||||  
QY 1671 GGTAGAAATATCAGCCCACTAATGCGCAAGTGAAGAAATTTGATTCAGCGCAGCATCC 1730  
QY 1081 CATGTTTGCAGCGAAAGCTTGGCTCGGGCGTTCGATCGTCCCAACCAAGCGGCGATTT 1140  
Db |||||  
QY 1731 TATTTTTCATCAGGTGTCTATGGGCAAGGTCTAGTTCATTTGAACCAAGCAAGGTGAGTT 1790  
QY 1141 AGTTTCTCGGTGAGTGAAGATTTGGTGGCATTTCCCATCTGGCCATGCTTTTCGCAAT 1200  
Db |||||  
QY 1791 GACCTCTCCAGTTAATGGACAGTACGCGTCTTTTCCCTACCAAGCATGCCATCGCAT 1850  
QY 1201 TCGCACCAAGGCTGAGGATGTTCCAAATGTGATATCTTGATGACATGTTGTTGACAC 1260  
Db |||||  
QY 1851 TGTCTCTGACGAGGAGTTGAAT-----TGCTCATCCATCGGTATGGATAC 1898  
QY 1261 AGTAAACCTCAACGGCAGCACTTTAAACCGCTGAAGAAGCAGGCGGATGAAGTCAAAGC 1320  
Db |||||  
QY 1899 AGTAGGCTTGTATGCAAGGTTTGAAGCTTTGTAGTCCAGAGATCAGTTACAGT 1958  
QY 1321 AGGGAGCTGCTGTGTAATTCGATTTGATGCCATTAAGGCTGAGGTTATGAGGTAAAC 1380  
Db |||||  
QY 1959 TGGTCAGCAACTGATTCGTTTGTATGATGATGATCATTAAGGCTGAGGTTGTTGACAGA 2018  
QY 1381 CACGCCGATTTGTTTTCGAA 1401  
Db |||||  
QY 2019 AACTCTGTTATCATCAAA 2039

RESULT 11

US-08-961-527-201/c

; Sequence 201, Application US/08961527

; Publication No. US20020032323A1

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 201:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3895 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; US-08-961-527-201

Query Match

Best Local Similarity 8.3%; Score 126.6; DB 8; Length 3895;

Matches 575; Conservative 0; Mismatches 619; Indels 27; Gaps 5;

QY 190 TTTAGATGTTGCCAAGCGGTTACAGGCAACCGCTTCTGCAAGCGACTCAAGGGCACTGCAGACTTCCT 249  
Db |||||  
QY 3218 TTTTACTGTTAATCGTATCGGTTACCAAGCCCAAGTTATCCAGCCTTGTTCGAGGTTT 3159  
Db |||||  
QY 250 GATTCGGCAACGATCGAAGTTCCTGCAAGCGACTCAAGGGCACTGCAGACTTCCT 309  
Db |||||  
QY 3158 GAGTCGTCTTATCTTGAATCTTCTGGCAACAGCATATCCAGAGTCATTTCTATGAT 3099  
QY 310 GATCACTCAGTGTGAGTGTCTGCTCACCGGATTCCTTACATTCATCGCAATTTGGCCC 369  
Db |||||  
QY 3098 TTTTGTACCTTTCTTGTCAATTCGATTCAGCCCTTGATTTTGGCTCATCTGTTTGGACC 3039  
QY 370 AGCAATGCGTGGGTGGCGATGTCTCGGCTCTGCTTCTGCTTACTCACCAGTCTCATCACTGG 426  
Db |||||  
QY 3038 AATCGTTGGCAATTTGGCAAGGACTTTTCATCAGTTGTCTTGGCAGGTTTAACTGGTCC 2979  
QY 427 CGGTGGTCCAGTCCGGGCTGCTCTTCTCGGCTCTGCTTACTCACCAGTCTCATCACTGG 486  
Db |||||  
QY 2978 AGTTAAATGGCTCTTCGCTGCAATTTTGGCGCCCTCTACGCTCAATTTGATCAACAGG 2919  
QY 487 TCTGACCACTGCTTCCCGCAATTTAGCTGGAGCTGTTT---AACAGGGTGGATCCTT 543  
Db |||||  
QY 2918 TCTGACCATATGACCAATGCAATTCATGATGATTCGCGATGCTGTTGGCACTGC 2859  
QY 544 CATCTTCGCAACGGCATCTATGCTTATATCGCCACGGGTGGCGCATGTTTGGCAGTGT 603  
Db |||||  
QY 2858 CCTATGGCCAAATGATGCTCTTCTTAAATATGCTCAAGGCTCAGCGGTGTTTGGCTATTA 2799  
QY 604 CTTCTCGCGCAAGAGTGAAGAGCTCAAGGGCC---TTGCAAGTCTTTCAGGTGCTCCGC 660  
Db |||||





```
1106 TCAGGCGTTCCTTCGATCCCAACCAAGGGGAGTTAGTTCTTCGCGTGAAGAAAT 1165
1519 CAAGGTATCGCTGTGAACCAAGTCAAGGGGTGCTCTATGCACTAGCTGATGCTGAAGTT 1578
1166 GTGGTGCATTCCTCATCTGCGCATCTTCGAGTTCGACGTCGACCAAGGTGAGGATGCTCC 1225
1579 TCAATGCCCTTCCAAAGGGGACGCTTTGGTTGAAACAAGAAATGG----- 1628
1226 AATGTGATATCTTCGATGCAATTTGGTTTCGACAGTAAACCTCAACGGCACGCACTTT 1285
1629 --TCTGAAGTTTGAATTCAGTTTGGTATTGACACTGTATCTATGAACGGTGAAGGTTT 1686
1286 AACCCGCTGAAGAGCAGGCGATGAAGTCAAGCAGGAGAGTCTGTGTGAATTCGAT 1345
1687 GAACCAAAAGTTGCTCAAGGTAAATAAGGTGAAGCTGGCGATGTTCTTTGGAACATTTGAC 1746
1346 ATTGATGCCATTAAGGCTGCAGGTTATGAGGTAAACCAAGCGATTTGTTTTCGAAT 1402
1747 TCAAAACAAATCGCTGCAGCTGGACTTGATGATACAAATGGTTATCGTTACAAAT 1803

RESULT 14
US-10-474-776-152
; Sequence 152, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYPEPTIDES
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 152
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-152

Query Match 7.1%; Score 107.8; DB 17; Length 1185;
Best Local Similarity 46.6%; Pred. No. 1.5e-22;
Matches 552; Conservative 0; Mismatches 597; Indels 36; Gaps 5;

302 GACTTCTTCATCATCTCCAGTGTGAGTGTGCTGCTCACCGGATTCCTTACATTCATCGCC 361
34 GACCTCTTGGTAAACCACTTCGTGACACTTTTGGTCACTGTCTATCTCTGGACTCTTTGTC 93
362 ATTGGCCCAAGCAATGCGTGGTGGCGATGTGCTGGCACAGTCTACAGGAGCTTTAT 421
94 ATTGACCAAGTTTCCAGCTTTGAAACTATCATCTATTTGCTACAAAGGATCTTT 153
422 GATTTTCGGTGTCCAGTGTGCGGCTGTGCTCTTCGGTCTGCTACTACCAATCGTCATC 481
154 ACATGCCAATTTGGTCTTGGTGTCTTGAATGTGGGGTTCACCAATGTATGTCGTG 213
482 ACTGTTCTGACAGTCTTCCTCCGCAATGTAGCTGGAGCTGTTHAACAGG---TGA 538
214 TCAGGTGTGCACACATCTTCAACTTGTGTAAGTGAATTTACTTCTGCTGCACCATGCT 273
539 TCTTCATCTTCGCAAGCACTATGCTATGCTATATCGCCAGGTCGCGCATGTTTGGCA 598
274 AACCCATTCAGCTATCATCAGCTGTGTATGACAGCTCAAGTGTGCTACTGTGCG 333
599 GTGTTCTTCCTGCGCAAGAGTGAAGAGCTCAAGGGCTTTCAGGTCCTTCAGGTGCTCC 658
334 GTTGTGTGTTTAAACAAATAATCCAAACTGAAACACTTGTCTTCCCGGCTGCTCTTCT 393
659 GCTGTTCTTGTATACGAGGCTTGCATCTTCGCTGTGAACCTTCCGCTGCGCTGGCGG 718
394 GCCTTCTAGTATTACAGAGCCTGCTATCTTCGGGGTGAATTTGGCTTCCGTAACCA 453
```

## RESULT 15

```
US-08-961-527-30/c
; Sequence 30, Application US/08961527
; Publication No. US2002003232A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
```

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/961,527  
;; FILING DATE:  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Brookes, A. Anders  
;; REGISTRATION NUMBER: 36,373  
;; REFERENCE/DOCKET NUMBER: PB340P1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (301) 309-8504  
;; TELEFAX: (301) 309-8512  
;; INFORMATION FOR SEQ ID NO: 30:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9769 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
US-08-961-527-30

Query Match 7.1%; Score 107.8; DB 8; Length 9769;  
Best Local Similarity 46.6%; Pred. No. 4.1e-22;  
Matches 552; Conservative 0; Mismatches 597; Indels 36; Gaps 5;

QY 302 GACTTCCTGATACCTCCAGTGTGAGCTTCTGCTACCGGATTCCTTCAATTCATCGCC 361  
DB |||||  
QY 8979 GACCTCTGGTAACACCAATTCGTGACACTTTTGGTCAATGCTATCTCTTGGACTCTTGTG 8920  
DB |||||  
QY 362 ATTGGCCACGATCGCTGGTGGGATGCTGTCGACACCGCTTACAGGACTTTTAT 421  
DB |||||  
QY 8919 ATTGGACCAAGTTTCCAGCTGTTGAAACCTACATCTTATTGCTACAAAGCAATCTT 8860  
DB |||||  
QY 422 GATTTCGTGGTCCAGTCCGGCGCTGCTCTTCGCTCTGCTCTACACCAATTCGTCATC 481  
DB |||||  
QY 8859 AGCATGCCATTTGGTCTTGGTGGTTTCTGATTTGGTGGGTTCCACCAATTTGATGCTG 8800  
DB |||||  
QY 482 ACTGTCGTGACAGTCTTCCGCCAATTTGAGCTGGAGCTGTTTAAACAGGG- ---TGGA 538  
DB |||||  
QY 8799 TCAGGTGTGCACCAATCTTCAACTTGTGTTGAATGCAATTTCTGCTGCTGACCAATGCT 8740  
DB |||||  
QY 539 TCCTTCATCTTCGACGCACTATGCTTAATATCGCCAGGTCGGCATGTTTGGCA 598  
DB |||||  
QY 8739 AACCCATTCACGCTATCATCAAGCTGCTATGACAGCTCAAGTGTCTGCTACTGTTGG 8680  
DB |||||  
QY 599 GTGTTCTTCCTGGCGAAGAGTGAAAGCTCAAGGCTTTCAGGCTGCTTCAGGTGCTTCC 558  
DB |||||  
QY 8679 GTTGGTGTAAACAAAATCCAAACTGAAAACACTTGTCTTCCCGCTGCTCTTTCT 8620  
DB |||||  
QY 659 GCTGTTCTTGGTATTAAGGAGCTGCGATCTTCGTTGTAACCTTCGCTGCGCTGGCGG 718  
DB |||||  
QY 8619 GCCTTCCTAGGTATTACAGAGCTGCTATCTTCGGGGTGAACCTTGGCTTCGGTAAACCA 8560  
DB |||||  
QY 719 TTCTTCATCGGTATCGGTACCGAGTATCGGTGGGCTTTGATTTGCACTCTTTAATATC 778  
DB |||||  
QY 8559 TTCTTCCTTCAATGATTGCTGCTGCTCAATCGGTGGTGAATGGCTTCTATCTCTTGACTT 8500  
DB |||||  
QY 779 AAGGCAAGTTGCTTGGCGCTGCTGAGTCTTTCGTTGGGTGTTGTTCTTATTTGATCTCCAGAT 838  
DB |||||  
QY 8499 GCTGTTACTGGTAATGTTATCACCATCATCTCCCTGGTACAACTCTTATGTTGGTAACGA 8440  
DB |||||  
QY 839 ATGTCATGTTCTTGGTGTGTCAGTTGTTTACCTTCTTCATCGCATTCGGCGCAGCGATT 898  
DB |||||  
QY 8439 CAACATCCACAAATACCTTCTTATGTTAGTGTATCATTT-----TGCCCTT 8395  
DB |||||  
QY 899 GCTTATGCTTTTACTTGGTTCGCGCAACGCGCAGCATTTGATCCAGATGCAACCGCTGCT 958  
DB |||||  
QY 8394 GGTTTGCTCTTACTTACATGTTTGGTTACG---AAGATGAAGTAGAGCAACTGCACT 8338  
DB |||||  
QY 959 CCAGTGCCTGAGGAACGACCAAGCCGAGCAAGCAACCGCAGCAATTTTCAACCAT 1018  
DB |||||  
QY 8337 GCATAACGAGCTGAAGTGGTGAAGAAAAGAAAGTGTGGCCAGCAGCTCTTCAAAAT 8278  
DB |||||

Search completed: October 30, 2004, 18:51:03  
Job time : 760 secs

QY 1019 TCACCATCATCCAGGACACCTTTGACCGGTGAAGTATTGCACTGAGAGCGTCAAGCAT 1078  
DB |||||  
QY 8277 GAAACACTTGT---AACTCCATATCGTCGGTGAATGTTGTCGCTCTTGTGATGTCGAATGAC 8221  
DB |||||  
QY 1079 GCCATGTTTCCAGCGGAAAGCTTGGCTCGGGCGTTGCCATCGTCCCAACCAAGGGGCGAG 1138  
DB |||||  
QY 8220 CCAGTCTTCTCAAGTGGAGCTATGGGAAGGATTCGTTGTGAACCAAGCAAGCGGCGTG 8161  
DB |||||  
QY 1139 TTAGTTTCTCGGTGAGTGGAAAGATTGTGGTGGCAATTCCTATCTGGCCATGCTTTCGCA 1198  
DB |||||  
QY 8160 GTCTATGCACGAGCTGATGCTGAAGTTTCAATTGCCCTTTTCCAACAGGGCAGCTTTTGGT 8101  
DB |||||  
QY 1199 GTTCGCACCAAGCTGAGGATGTTTCCAAATGTGGATATCTTGAAGCAATTTGGTTTCGAC 1258  
DB |||||  
QY 8100 TTGAAAACAAGAAATGGTGCT-----GAAGTTTGTGATTCATGTTGTGTTATTTGAT 8053  
DB |||||  
QY 1259 ACAGTAAACCTCAACGGCACCGCACTTTTAAACCGCTGAAGAGCAGGGCGATGAAGTCAAA 1318  
DB |||||  
QY 8052 ACTGTATCTATGAACGGTGAACGGTTTGAACAAAAGTTGCTCAAGGTAATAAGGTGAAA 7993  
DB |||||  
QY 1319 GCAGGGAGCTGCTGTGTGAATTCGATATTTGATGCCATTAAGGCTGCAGGTTATGAGGTA 1378  
DB |||||  
QY 7992 GCTGGCGATCTTCTTGGAACATTTGACTCAACAAAATCGCTGCAGCTGCACTTGATGAT 7933  
DB |||||  
QY 1379 ACCACGCCGATTTGTTTTCGAAATTAAGAAACCGGACCTGTAAACACTTACGTTTG 1438  
DB |||||  
QY 7932 ACAAATGTTTATCGTTTACAAATACAGGTGACTACGCTTCAGTAGCTCCAGTCGCAACA 7873  
DB |||||  
QY 1439 GCGAAATTAAGCGGGAGCCAACTGCTCAAGCTCGCAAGAAA 1483  
DB |||||  
QY 7872 GGTTCAGTTGCTAAGGGGATGCTGTGATCGAAGTGAATAATCTAA 7828  
DB |||||

**This Page Blank (uspto)**

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 13:44:46 ; Search time 4437 Seconds  
(without alignments)  
12540.787 Million cell updates/sec

Title: US-09-604-231-1  
Perfect score: 1527  
Sequence: 1 ctcattggctctgcgcgtt.....gttgaaccttgagtgttcg 1527

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gssi:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	105.4	6.9	917	8	BZ685850
C 2	92.6	6.1	750	9	CL686986
C 3	80.8	5.3	770	7	CN823265
C 4	73.2	4.8	861	9	CL657583
C 5	63.2	4.1	1412	9	CG744886
C 6	62	4.1	788	9	CL693618
C 7	48.8	3.2	680	1	AL506262
C 8	49.8	3.2	775	9	CL675911
C 9	47.4	3.1	1328	8	BZ572810
C 10	46.4	3.0	763	9	CL695497
C 11	45.8	3.0	473	7	TO3017
C 12	45.4	3.0	694	4	EM100823
C 13	45.4	3.0	791	9	CL661694
C 14	44.4	2.9	785	6	CD443326
C 15	43.2	2.8	869	7	CK159167
C 16	42.6	2.8	473	9	CE077138
C 17	42.2	2.8	798	6	CD780937
C 18	42	2.8	750	4	BM559084
C 19	41.8	2.7	681	9	CNS02EOD
C 20	41.8	2.7	889	9	CK159613
C 21	41.6	2.7	895	9	CNS0071A
C 22	41.4	2.7	477	2	BE997958
C 23	41.4	2.7	518	6	CA920257
C 24	41.4	2.7	547	6	CD442889

25	41.4	2.7	559	2	AW574083	AW574083	EST316674
26	41.4	2.7	601	4	BG580528	BG580528	EST482254
27	41.4	2.7	627	2	BE998413	BE998413	EST430136
28	41.4	2.7	793	6	CF069882	CF069882	EST670603
29	41.2	2.7	511	5	BU345319	BU345319	604169160
C 30	41	2.7	581	5	BQ988486	BQ988486	QGF15A22.
C 31	41	2.7	683	5	BQ986937	BQ986937	QGF10M14.
C 32	41	2.7	732	5	BQ933574	BQ933574	QGF4e10.Y
C 33	41	2.7	739	5	BQ991717	BQ991717	QGF23J24.
C 34	41	2.7	744	5	BQ988795	BQ988795	QGF15003.
C 35	41	2.7	745	5	BQ986791	BQ986791	QGF10H09.
C 36	41	2.7	759	5	BQ993437	BQ993437	QGF3a03.Y
C 37	41	2.7	776	5	BQ995038	BQ995038	QGF8003.Y
C 38	40.8	2.7	884	9	CNS006UO	CNS006UO	AL065923
C 39	40.8	2.7	983	9	CNS03CVD	CNS03CVD	Drosophil
C 40	40.8	2.7	1101	9	CNS00LXJ	CNS00LXJ	AL078805
C 41	40.2	2.6	970	9	CNS03H6V	CNS03H6V	Tetraodon
C 42	40	2.6	902	8	CC366418	CC366418	AL243904
C 43	39.8	2.6	548	7	CK895185	CK895185	PUHEW79TB
C 44	39.8	2.6	606	2	BE997957	BE997957	SGP154437
C 45	39.6	2.6	714	7	CF753589	CF753589	EST429680

## ALIGNMENTS

RESULT 1  
BZ685850/c  
LOCUS PUB666TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa035K11,  
DEFINITION genomic survey sequence.  
ACCESSION BZ685850  
VERSION BZ685850.1 GI:28245462  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 917)  
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.  
Maize Genomics Consortium  
Unpublished (2003)  
CONTACT: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@cigr.org  
Seq primer: 1P  
Class: sheared ends.  
Location/Qualifiers  
1. 917  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBTa035K11"  
/clone\_lib="ZM 0.6 1.0 KB"  
/note="Vector: pCR4-TOPO; Site:1: EcoRI; 0.6-1.0 kb high  
CoT selected genomic DNA library"

Query Match 6.9%; Score 105.4; DB 8; Length 917;  
Best Local Similarity 52.6%; Pred. No. 1.2e-18;  
Matches 308; Conservative 0; Mismatches 261; Indels 16; Gaps 3;  
Qy 1 CTATGTCATCTGCGCGCTTGGCTTTCGACGTGTTGGTTCACCGCACCAAG 60  
Db 585 CTATGTCACCAACACCTCGTTCGTTCTCTCCACGCGTGGTGGCTGGTGGCCCAAG 526

QY	61	CGTTTCGGCGGCAATGAGTTCCTGGGCGCGGTATTTGATGCGGATGCTGTTCCGAG	120
Db	525	CGCTTTGGCGGCAAGTAAATTCCTCGGATCGTG-CITGGCCTGATCTGTGTGATCAGA	467
QY	121	CTTGTGTAACGGCTACAGCTGGCGCGCCACCATCGGTGCGGGGGAATGCAATGTGTC	180
Db	466	CTGTCTCAACCCCTGGAATACCGCAAGCGGTGCGCGGTCTCGACGGCCAGAGCTGCC	407
QY	181	CTGTGTTGGT-----TTAGATGTGCCCCAAGCGGTTTACAGGACCCGCT	228
Db	406	GTACTTCGATATTTTCGTTTGGTTCAAGATTTGAGAAAGTGGGTACACAGGGCGAGATCCT	347
QY	229	TCCTGTCTGGTGTTCCTGGAATCTTGCAACGATCGAAGTTCCTGCAACAGCACT	288
Db	346	GCCGATCCTGATGCGCCCTATGTGATGAGCGTGATCGAAGAAATGGCTGCGGGCGGAGT	287
QY	289	CAAGGGCACTGCACATTCCTGATCACTCCAGTCTGCTGACGTTGCTGCTCACCGGATTCCT	348
Db	286	GCCCAACGGATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	227
QY	349	TACATTCATCCCATTCGCGCCAGCAATGCGCTGGTGGGCGATGTGCTGGACACGCTCT	408
Db	226	CGCCCTGCGCAATATCGCGCCGCTGACCCGCTCACTGGGCACTCTGATCACCGAAGCGT	167
QY	409	ACAGGACTTTATGATTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	468
Db	166	GGTCACCTTGTGACCTCGGACCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	107
QY	469	ACCAATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	525
Db	106	GCGCTGTGATACCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	47
QY	526	TAAACAGGCTGATCTTCATCTTCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	570
Db	46	TACTCAGGGCGCACCTTCATCTGCGCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT	2
RESULT 2			
LOCUS	CL686986	750 bp	DNA
DEFINITION	PR10145c.F12.2 - PR10145c.BR (750) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.		
ACCESSION	CL686986	1	GI:50195625
VERSION	GSS.		
KEYWORDS	Pristionchus pacificus		
SOURCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.		
ORGANISM	1 (bases 1 to 750)		
REFERENCE	Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.		
AUTHORS	AppADB: an AcedB database for the nematode satellite organism		
TITLE	Pristionchus pacificus		
JOURNAL	Nucleic Acids Res. 32 (1), D421-D422 (2004)		
COMMENT	Contact: Sommer RJ Evolutionary Biology Max-Planck-Institute for Developmental Biology Spemannstr. 37-39, Tuebingen D-72076, Germany Tel: 00497071601371 Fax: 00497071601498 Email: ralf.sommer@tuebingen.mpg.de This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada. Seq primer: T7 Class: fosmid ends.		
FEATURES			
source	1. .750		
	/organism="Pristionchus pacificus"		
	/mol_type="genomic DNA"		
	/strain="California"		
	/db_xref="taxon:54126"		
	/clone_lib="Mixed stage fosmid library of P. pacificus"		
ORIGIN			
Query Match	6.1%;	Score 92.6;	DB 9; Length 750;
Best Local Similarity	53.9%;	Pred. No. 5.5e-15;	
Matches 241;	Conservative	0; Mismatches 194;	Indels 12; Gaps 2;
QY	1028	ATCCAGGCACTTTGACCGGTGAAGCTATTGCACTGAGCAGCTCAGCAGTGCATGTTT	1087
Db	21	ATCTGTTCCAGATGACGGAGAGATTGTCCGCTCATTCACGTCGCTGATACACGTTT	80
QY	1088	GCCAGCGGAAAGTTGGCTCGGCGTTGCCATCGTCCCAACCAAGCGGCACTGTTTCT	1147
Db	81	GCCAGTGGCTGTGGGTAAGGTAICGCCATTTCGCCCTCGGTTGGTGAAGTCGTTCT	140
QY	1148	CCGCTGAGTGAAGATTGTGGCATTCGCCATTCGCCATTCGCCATTCGCCATTCGCCAC	1207
Db	141	CCGCTTCGGGTGCAATTGCTTGTGTTTCCCAATTACA-----CGCCATTGGCATT	194
QY	1208	AAGCTCAGGATGTTTCCAATGTGGATATCTTGTGACACATTTGTTTCGACACAGTAAAC	1267
Db	195	GAGTCAGATGATGG-----TGTGAGATCTCTGATTCATGTCGTAACACCGTAAAA	248
QY	1268	CTCAACGGCAACCTTTAAACCCCTGAAGAGCAGGGCGATGAAGTCAAGCAGGGGAG	1327
Db	249	CTGACGGCAATTCCTTTTCGCTCAGTCAACGTCGAGTGAACAGGTCATACAGGCGAT	308
QY	1328	CTGCTGTGATTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1387
Db	309	CGGCTGATTTCTTTGATATCCCTGCTATTCGCGAGCGCGGATTTGATGTGACGCGCG	368
QY	1388	ATTGTTGTTTCGAATTCAGAGAAACCGGACCTGTAAACATCTACGTTTGGGCGAATT	1447
Db	369	GTATTATCACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	428
QY	1448	GAAGCGGAGCAACCTGCTCAACGTC	1474
Db	429	AGCGCAGGTGAACCGCTGTTATCCATC	455
RESULT 3			
LOCUS	CN823265	770 bp	mRNA
DEFINITION	Oa_splbn_04M09_M13reverse Sheep spleen\brain spsort1 library Ovis aries cDNA clone Oa_splbn_04M09 5', mRNA sequence.		
ACCESSION	CN823265	1	GI:47951334
VERSION	EST.		
KEYWORDS	Ovis aries (sheep)		
SOURCE	Ovis aries		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis. 1 (bases 1 to 770)		
REFERENCE	Gossner,A. and Hopkins,J.		
AUTHORS	Ovine spleen\brain cDNA library		
TITLE	Unpublished (2004)		
JOURNAL	Contact: J Hopkins		
COMMENT	Veterinary Biomedical Sciences University of Edinburgh Summerhall Square, Edinburgh, EH9 1QH. Email: j.hopkins@ed.ac.uk Plate: 04 row: M column: 09 Seq primer: M13reverse High quality sequence start: 15 High quality sequence stop: 547. Location/Qualifiers		
FEATURES			
source	1. .770		
	/organism="Ovis aries"		
	/mol_type="mRNA"		
	/db_xref="taxon:9940"		
	/clone="Oa_splbn_04M09"		



vector."

ORIGIN

Query Match 4.1%; Score 63.2; DB 9; Length 1412;  
Best Local Similarity 46.8%; Pred. No. 2.1e-06;  
Matches 274; Conservative 0; Mismatches 308; Indels 6; Gaps 2;  
172 AATGTGTCCTCTGTTGGTTAGATGTTGCGCAAGCGGTTACAGGCGACCGTCTTC 231  
Db |||||  
655 AGTGTGGCATTTGGCATGTTTCAGCATCGGCAAGTGGCTATCAGGCACAGGTGATCCC 596  
Qy |||||  
232 TGTGCTGTTGTTCTTGTGATTTGCGCAATCGAGAGTCTCTGCAACAGGACTCAA 291  
Db |||||  
595 GGCACCTTTAGCCGGGCTGGCAGCTGGGCGTTATTGAACTCGCCTTAAACATCGTACC 536  
Qy |||||  
292 GGGCACTGCAGACTTCTGATCACTCCAGTCTGACGTTGCTGTCTACCGGATTCCTTAC 351  
Db |||||  
535 GGATTACCTCTATCTGTTGGTGGTGCCGCTGTTGCTGATCTCGCGGTGTTCTCGC 476  
Qy |||||  
352 ATTCAATCGCCATGCGCCAGCAATGCGCTGGTGGGCGATGCTGCGCACACGCTCTA-- 409  
Db |||||  
475 CCATGCGCTGATTGGTTCGGTTTGGTTCGCATGATTGGCGATGGCGTTCGCTTTCGGTACG 416  
Qy |||||  
410 -CAGGACTTTATGATTTCCGTTGTCAGTCCGCGCTCTGCTTTCGCTCTGTTCTACTC 468  
Db |||||  
415 TCACCTGATGACCGGCGAGCTTTGCTCCGATTGGTGGCGCATGTTTGGCTCTCTGTAGCG 356  
Qy |||||  
469 ACCAATGCTCATCTGCTGTCACACGAGTCTCTCCCGCAATGAGCTGGAGCTGTTT-- 526  
Db |||||  
355 CCGCTGGTATCACCGCGTACACACGACACCCCTTCTATTGATTGCGAGATGATTCA 296  
Qy |||||  
527 -AACGAGGTGGATCCTTCACTTCGCAAGCGCATCTATGGCTATATATCGCCAGGTGC 585  
Db |||||  
295 AAGCATGGTGGCAGCCGAGTGGCGCTGATTGCGCTGTCGAATATGCTCAGGGCTC 236  
Qy |||||  
586 GGCATGTTTGCAGTGTCTTCTCGCGAAGTCAAAAGCTCAAGGCGCTTCGAGGTGC 645  
Db |||||  
235 CGCGTGATGAGCATATCAATTCAGCGCGGAGCAATGACATGACCGAGATCTCCGTGCC 176  
Qy |||||  
646 TTCAGTGTCTCCGCTGTTCTTTGGTATTACGAGCGCTCGATCTTCGTTGGAACCTCG 705  
Db |||||  
175 TGCCGCTATCTCCGCTGCTGGTCTACTGAGCTGCAATGTACGCGCATCAACCTGAA 116  
Qy |||||  
706 CTGCGCTGCGCTTCTTCATCGTATCGGTACCGCAGTATCGGTGG 753  
Db |||||  
115 ATATCGCTTCCGAGTCTGTGCGCATGATGTTCTTGTCTGCGCAGG 68

RESULT 6  
CL693618 788 bp DNA linear GSS 10-JUL-2004  
LOCUS PRI0162a\_D06\_2 - PRI0162a.BR (788) Mixed stage fosmid library of P.  
pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.  
CL693618  
CL693618.1 GI:50215526  
GSS.  
Pristionchus pacificus  
Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.  
1 (bases 1 to 788)  
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.  
AppaDB: an AcedB database for the nematode satellite organism  
Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de

FEATURES

source

Location/Qualifiers  
1..788  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage fosmid library of P. pacificus  
var. California"  
/note="Vector: pPifos-5 Fosmid vector"

ORIGIN

Query Match 4.1%; Score 62; DB 9; Length 788;  
Best Local Similarity 52.3%; Pred. No. 3.8e-06;  
Matches 137; Conservative 0; Mismatches 125; Indels 0; Gaps 0;  
1229 GTGGATATCTTGATGCACTTGGTTTCGACACATTAACCTCAACGGCAGCATTAAAC 1288  
Db |||||  
388 GCGGAGATCGTCGTCATATGGGTATCGACACCGTAGCGTGAAGGTAAAGGCTTTAAA 447  
Qy |||||  
1289 CCGCTGAAGAGCGGGCGATGAAGTCAAGCAGGAGCTGCTGTGTAATTCGATATT 1348  
Db |||||  
448 CCGTCTGTGAAGAGGGGGCGCAGGTAAGCGCAGGCAACCGATTCTGGAATGATCTG 507  
Qy |||||  
1349 GATGCCATTAAAGCTGCGAGGTTATGAGTAACACCGCGATTGTTTTCGAATTACAAG 1408  
Db |||||  
508 GATTACCTGAACGCTTAACGCCCGCTCGATGATTAGCCGGTGGTTTCGAGCAATATCGAC 567  
Qy |||||  
1409 AAAACCGGACCTGTAACACTTACCGTTTGGCGCAATTAAGCGGAGCAGCACTGCTC 1468  
Db |||||  
568 GATTTCAGTGCCTTGATCATTAAAGCTCAGGCGCATGTTGTGCGCGTCAACACCGCTG 627  
Qy |||||  
1469 AACGTCGCAAAAGAGCGG 1490  
Db |||||  
628 TATGAATCAAAAAGTAATCTG 649

RESULT 7

AL506262/c

LOCUS

DEFINITION

AL506262 Hordeum vulgare Barke developing caryopsis (3.-15.DAP)

Hordeum vulgare subsp. vulgare cDNA clone HY02124T 5', mRNA

sequence.

AL506262

AL506262.1 GI:12032477

EST.

KEYWORDS

SOURCE

ORGANISM

Hordeum vulgare subsp. vulgare

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poaceae; Poaceae; Hordeum.

1 (bases 1 to 680)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Michalek W

Institute for Plant Genetics and Crop Plant Research

Corrensstr.3, D-06466 Gatersleben, Germany

Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de

Seq primer: T3 primer for 5'end.

Location/Qualifiers

1..680

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="barke"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/clone="HY02124T"

/tissue\_type="developing caryopsis (3.-15.DAP)"

/lab\_host="XL0LR"





QY 755 GC 756  
Db 541 GC 542

RESULT 11  
T03017/c  
LOCUS  
DEFINITION FB20C8 Fetal brain, Strata gene Homo sapiens cDNA clone FB20C8, mRNA  
ACCESSION T03017  
VERSION T03017.1 GI:314258  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 473)  
Khan, A.S., Wilcox, A.S., Polymeropoulos, M.H., Hopkins, J.A., Stevens, T.J., Robinson, M., Orpana, A.K. and Sikela, J.M.  
TITLE Single pass sequencing and physical and genetic mapping of human CDNA  
JOURNAL Nat. Genet. 2, 180-185 (1992)  
MEDLINE 94258200  
PUBMED 1345165  
COMMENT On Sep 21, 1992 this sequence version replaced gi:279156.  
Contact: Sikela JM  
Department of Pharmacology  
University of Colorado Health Sciences Center  
Box C236, 4200 E. 9th Ave, Denver CO 80262-0236  
Tel: 3032708637  
Fax: 3032707097  
Email: nikki@tally.uchsc.edu.

FEATURES  
source  
1..473  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):86662"  
/db\_xref="GDB:D0S2785E"  
/db\_xref="taxon:9606"  
/clone="FB20C8"  
/clone\_lib="Fetal brain, Strata gene"  
/note="Vector: Lambda ZAP II; The FB library (catalog #937201, Strata gene) was constructed by directional cloning and oligo d(T)-priming in Lambda ZAP II phage utilizing mRNA from a 17 week human fetus (total brain)."

ORIGIN  
Query Match 3.0%; Score 45.8; DB 7; Length 473;  
Best Local Similarity 59.7%; Pred. No. 0.15;  
Matches 77; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1231 GGATATCTTGATGCAATGTTTCGACACAGTAAACCTCAACGCGCAGCCTTTAAACC 1290  
Db 275 GGAGATGTCGTCCATATGGGTATCGACACCGTAGCGTGGAGTAAAGCGCTTTAAACG 216

QY 1291 CTGAGAGACGAGCGCGATGAAGTCAAGCAGGAGCGTCTGTGTGAATTCGATATGA 1350  
Db 215 TCTGGTGAAGAGGGTGGCGAGGTAAAGCGCAGGCAACCGATCTCGAATCGATCTGGA 156

QY 1351 TGCCATTAA 1359  
Db 155 TTACCTGAA 147

RESULT 12  
BM100823/c  
LOCUS  
DEFINITION BM100823 694 bp mRNA linear EST 23-JUL-2002  
Ebp101\_SQ002\_B17 R pistil, 1 DPA, no treatment, cv Optic, Ebp101  
Hordeum vulgare subsp. vulgare cDNA clone Ebp101\_SQ002\_B17 5', mRNA  
sequence.  
ACCESSION BM100823  
VERSION BM100823.2 GI:21943888

KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.  
1 (bases 1 to 694)  
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.  
TITLE Development of Barley Transcriptome Resources  
JOURNAL Unpublished (2001)  
COMMENT On Nov 21, 2001 this sequence version replaced gi:17031891.  
Contact: Waugh R, Marshall DF  
Genome Dynamics/Computational Biology  
Scottish Crop Research Institute  
Invergowrie, Dundee, DD2 5DA, Scotland, UK  
Tel: 00 44 1382 562731  
Fax: 00 44 1382 562426  
Email: est@scri.sari.ac.uk  
All sequence has a Phred quality score of 20 or over  
Seq primer: M13 reverse.

FEATURES  
source  
1..694  
Location/Qualifiers  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="Optic"  
/sub\_species="vulgare"  
/db\_xref="taxon:112509"  
/clone="EBp101\_SQ002\_B17"  
/tissue\_type="pistil"  
/dev\_stage="1 DPA"  
/lab\_host="DH10B"  
/clone\_lib="pistil, 1 DPA, no treatment, cv Optic, Ebp101"  
/note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I; Non-normalised library, directionally cloned into pSPORT1. Derived from pistils dissected from developing grains (24 hours post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."

ORIGIN  
Query Match 3.0%; Score 45.4; DB 4; Length 694;  
Best Local Similarity 48.6%; Pred. No. 0.22;  
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 636 TTGCAGGTGCTTCAGGTCTCCGCTGTTCTGGTATTACGAGCCTCGATCTCGGTG 695  
Db 693 TTGTTGTTGTTCTTCGTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 634

QY 696 TCACCTTCGCTGCGCTGGCGGTTCTTCATCGGTATCGGTACCGACATCGGTGGCG 755  
Db 633 TCATTCGTATTGTTTTCGCCAGTTTCGTCGTTGTTGTTGTTGTTGTTGTTGTCG 574

QY 756 CTTTGATTGCACTCTTTAATATCAAGCAGTTCGTTGGCGCTGCAGGTTCCTGGGTG 815  
Db 573 TTGCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 514

QY 816 TTGTTTCATTGATGCTCCAGATATGGTCATGTTCTTGGTGTGTCAGTTGTTACCTTCT 875  
Db 513 TTGTTATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 875

QY 876 TCATCGCATTCGGCG 890  
Db 453 TCATTGTTATCGTCG 439

RESULT 13  
CL661694/c  
LOCUS  
DEFINITION CL661694 791 bp DNA linear GSS 09-JUL-2004  
PRI013c.B10 - PRI013c.B21 (791) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

REFERENCE AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 785) Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and Messing, J.	
TITLE JOURNAL COMMENT	Sequencing of the maize endosperm ESTs Unpublished (2002) Contact: Lai, Jinsheng Dr. Joachim Messing's lab Waksman Institute, Rutgers University 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA Tel: 732-445-3801 Fax: 732-445-5735 Email: jlai@waksman.rutgers.edu	
FEATURES	Seq primer: T3. Location/Qualifiers 1..785 /organism="Zea mays" /mol_type="mRNA" /cultivar="W22" /db_xref="taxon:4577" /tissue_type="Endosperm of 7-23DAP" /clone_lib="Endosperm_4" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"	
ORIGIN	2.9%; Score 44.4; DB 6; Length 785; Query Match Best Local Similarity 46.9%; Pred. No. 0.45; Matches 138; Conservative 0; Mismatches 156; Indels 0; Gaps 0;  QY 641 GGTGCTTTCAGGTCTCCGCTGTTCTTGTTATTAACGAGCCCTGCGATCTTCGGTGGAAAC 700 DB 457 GGTGTTGCTCATGGCTCTGTCGTGTGTTGCACATGGCGCTTGTCATTGTTGGTCATGTT 398 QY 701 CTTGCGCTCGCGTGGCGGTTCTTCATCGATACGCTACCGACGATATCGGTGCGCTTTG 760 DB 397 GTTGTGTTGGTGCTCTGTTGTGTGTTGTGTTGGTCTGCGGTTGTTGTGTTGTTGTT 338 QY 761 ATTGCACCTCTTTAATATCAAGGCAGTTCGCTTGGCGCTGCGAGTTTCTTGGGTGTTGTT 820 DB 337 GAACATGAACCTCTTGTGTTGCTGATGTTTGTGTTGCATGTCAACTGTTGTTGTTGTT 278 QY 821 TCTATTGATGCTCCAGATATGTCATGTTCTTGGTGTGTGCAGTTGTTTACCTTCATC 880 DB 277 GGTGTTGCTCTGCTGTGTGATGTGTGGTGGTGTTGATGTTGTTGTTTGGATGAT 218 QY 881 GCATTCGGCCACGATGCTGTTATGGCTTTACTTGCTTCGCGCAACGCGACG 934 DB 217 GTTGTGTTGGATGATGTTGCTGCTCAGGCTGTTGTTGACACACACCTGCACG 164	
RESULT 15	CK159167 869 bp mRNA linear EST 05-DEC-2003	
LOCUS	FGAS040564 Triticum aestivum FGAS: TaU5 Triticum aestivum cDNA,	
DEFINITION	mRNA sequence.	
ACCESSION	CK159167	
VERSION	CK159167.1 GI:38985053	
KEYWORDS	EST.	
SOURCE	Triticum aestivum (bread wheat)	
ORGANISM	Triticum aestivum	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.	
AUTHORS	1 (bases 1 to 869) Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D., Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A., Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D., Penniket, C., Roach, J.L. and Sarhan, F.	
TITLE	Functional Genomics of Abiotic Stress In Wheat and Canola Crops	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Wm L Crosby	

## Bioinformatics

University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769  
Fax: 306 966 2033

Email: fgas\_est@cs.usask.ca

This sequence is the direct result of the Base calling software  
Phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [128,636].

Plate: Talet537 row: N column: 23.

## FEATURES

## source

Location/Qualifiers

1..869

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Wheat line PI 178383"

/db\_xref="taxon:4565"

/lab\_host="DH5 alpha"

/clone\_lib="Triticum aestivum FGAS: Talet5"

/note="Organ: Crown; Vector: pGEM-T; SSH (suppression

subtractive hybridization) cDNA library from genotype

PI178383 cold hardened at 2 C for 21 days and 49 days

(equal amount of cDNA pooled together before subtraction,

tester) and subtracted against genotype Norstar cold

hardened at 2 C for 1 day (24 H) (driver). Modified Smart

cDNA (Clontech) priming and non-directional cloning"

## ORIGIN

```
Query Match          2.8%; Score 43.2; DB 7; Length 869;
Best Local Similarity 46.9%; Pred. No. 1;
Matches 135; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 632 GGCCTTGACGGTCTCAGGTGTCCTCGCTCTCTGTGTTATTACGAGCGCTGGATCTTC 591
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 GTCTGTGTCGTGTGGTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 389
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 692 GGTGTAACTTCGCTCGCTGGCGGTCTTCATCGGTATCGGTACCGCAGCTATCGGT 751
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 GTTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 449
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 752 GCGGCTTTGATTCGACTCTTTAATATCAAGGAGTTGCGTTGGCGCTGCAGTTTCTTG 811
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 450 GTTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 509
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 812 GGTGTGTTTCTATTCATGCTCCAGATATGTCATGTTCTTGGTGTGTCAGTTGTTACC 871
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 510 GTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 569
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 872 TTCCTCATCGCATTCGGCGCAGGATTCCTATGCGCTTTACTTGGTT 919
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 570 GTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 617
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 30, 2004, 17:04:56  
Job time : 4442 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 22:57:53 ; Search time 197 Seconds  
(without alignments)  
1366.880 Million cell updates/sec

Title: US-09-604-231-2  
Perfect score: 2363  
Sequence: 1 MAWFFSLVNGYDVATMAA.....IEAGANLLNVAKEAVPATP 468

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02 : \*  
1: uniprot\_prot.\*  
2: uniprot\_tmbl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2363	100.0	661	2 Q8NMD6	Q8nmd6 corynebacte
2	2363	100.0	661	2 CAF21304	Caf21304 corynebacte
3	1062.5	45.0	651	2 Q88ZV9	Q88zv9 lactobacill
4	1058.5	44.8	651	1 PTSA PEDPE	P43470 pediococcus
5	977	41.3	534	2 Q93ML1	Q93ml1 lactobacill
6	962	40.7	650	2 Q9S6S6	Q9s6s6 lactococcus
7	953.5	40.4	664	1 PTSA STRMU	P12655 streptococc
8	914.5	38.7	647	2 Q74HI8	Q74hi8 lactobacill
9	914.5	38.7	647	2 AAS09702	Aas09702 lactobaci
10	901	38.1	649	2 Q7WWP7	Q7wwp7 lactobacill
11	832.5	35.2	643	2 Q82YR5	Q82yr5 enterococcu
12	759.5	32.1	627	2 Q9L8G6	Q9l8g6 clostridium
13	715	30.3	626	2 Q931B4	Q931b4 enterococcu
14	701	29.7	480	1 PTSB STAXY	P51184 staphylococ
15	695.5	29.4	630	2 Q8KGI9	Q8kgi9 bacillus ha
16	695	29.4	617	2 Q8Y904	Q8y904 listeria mo
17	678	28.7	474	1 PTSB PASMU	Q9cj22 pasteurella
18	674	28.5	633	2 Q82F57	Q82f57 listeria in
19	673.5	28.5	636	1 PTBA_BACHD	Q9kf90 bacillus ha
20	673	28.5	635	2 Q97NW9	Q97nw9 streptococc
21	670.5	28.4	644	2 Q8DUF2	Q8duf2 streptococc
22	667	28.2	634	2 Q8YAT6	Q8yat6 listeria mo
23	667	28.2	634	2 Q725E2	Q725e2 listeria mo
24	667	28.2	634	2 AAT02819	Aat02819 listeria
25	664	28.1	481	2 Q8CN82	Q8cn82 staphylococ
26	664	28.1	627	2 Q97PB8	Q97pb8 streptococc
27	663.5	28.1	676	2 Q8E208	Q8e208 streptococc
28	660	27.9	627	2 Q8DNS8	Q8dns8 streptococc
29	659	27.9	479	2 Q9KLT8	Q9klt8 vibrio chol
30	659	27.9	480	2 Q6G6U2	Q6g6u2 staphylococ
31	659	27.9	480	2 Q8NV35	Q8nv35 staphylococ

32	659	27.9	480	2 Q99RQ0	Q99rq0 staphylococ
33	659	27.9	480	2 Q7A3V6	Q7a3v6 staphylococ
34	659	27.9	480	2 Q88SA6	Q88sa6 lactobacill
35	658.5	27.9	644	2 Q9KJ80	Q9kj80 streptococc
36	658	27.8	480	2 Q6GE54	Q6ge54 staphylococ
37	658	27.8	705	2 Q8DN16	Q8dni6 streptococc
38	657	27.8	609	1 PTBA_BACSU	P40739 bacillus su
39	656.5	27.8	676	2 Q8E7G5	Q8e7g5 streptococc
40	655	27.7	655	2 Q8KST1	Q8kst1 streptococc
41	654.5	27.7	674	2 Q79W20	Q79w20 streptococc
42	654.5	27.7	674	2 Q8K5K9	Q8k5k9 streptococc
43	652	27.6	627	2 Q8NZJ7	Q8nzj7 streptococc
44	648.5	27.4	624	2 Q88SB9	Q88sb9 lactobacill
45	648.5	27.4	674	2 Q8NZ40	Q8nz40 streptococc

## ALIGNMENTS

## RESULT 1

ID	Q8NMD6	PRELIMINARY;	PRT;	661 AA.
AC	Q8NMD6; Q6M2J8;			
DT	01-OCT-2002 (T-EMBLrel. 22, Created)			
DT	01-OCT-2002 (T-EMBLrel. 22, Last sequence update)			
DT	01-OCT-2004 (T-EMBLrel. 28, Last annotation update)			
DE	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific (EC 2.7.1.69) (ENZYME II SUCROSE PROTEIN).			
DE	Name=ptsS; OrderedLocusNames=Cgl2642, cg2925;			
GN	Corynebacterium glutamicum (Brevibacterium flavum).			
OS	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
OX	NCBI_TaxID=1718;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;			
RA	Nakagawa S.;			
RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;			
RX	PubMed=12948626;			
RA	Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,			
RA	Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,			
RA	Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,			
RA	McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puchler A.,			
RA	Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,			
RA	Tauch A.;			
RT	"The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins.";			
RT	J. Biotechnol. 104:5-25(2003).			
DR	EMBL; AP005282; BAC00036.1; -			
DR	EMBL; BX927155; CAF21304.1; -			
DR	HSSP; P45618; 2GPR.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0008982; F:protein-N(Pi)-phosphohistidine-sugar phospho. . ; IEA.			
DR	GO; GO:0005351; F:sugar porter activity; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . ; IEA.			
DR	GO; GO:0006810; P:transport; IEA.			
DR	InterPro; IPR011055; Dup_hybrid_motif.			
DR	InterPro; IPR001996; Ptrans_EIIB.			
DR	InterPro; IPR003352; Ptrans_EIIC.			
DR	InterPro; IPR01127; Pts_EIIA.			
DR	InterPro; IPR010973; Pts_EIIB_sucr.			
DR	Pfam; PF00358; Pts_EIIA_1; 1.			
DR	Pfam; PF00367; Pts_EIIB; 1.			
DR	Pfam; PF02378; Pts_EIIC; 1.			
DR	ProDom; PD001476; Ptrans_EIIB; 1.			
DR	ProDom; PD002243; Pts_EIIA; 1.			

DR TIGR00830; PTB8; 1.  
DR TIGR01996; PTS-II-BC-sucr; 1.  
DR PROSITE; PS01035; PTS\_EIIB\_CYS; 1.  
KW Complete proteome; Transferase.  
SQ SEQUENCE 661 AA; 69148 MW; 69148 MW; COE5E1C7833B9F74 CRC64;

Query Match 100.0%; Score 2363; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 3.6e-144;  
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVFPVLNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 60  
DB 194 MAMVFPVLNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 253  
QY 61 KRLKGTADFLITPVLITLLTGFLTFIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFLGL 120  
DB 254 KRLKGTADFLITPVLITLLTGFLTFIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFLGL 313  
QY 121 VYSPVITGLHOSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 180  
DB 314 VYSPVITGLHOSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 373  
QY 181 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 240  
DB 374 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 433  
QY 241 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEABAP 300  
DB 434 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEABAP 493  
QY 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIIVAFP 360  
DB 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIIVAFP 553  
QY 361 SGHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELICEFDIDA 420  
DB 554 SGHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELICEFDIDA 613  
QY 421 AAGVEVTPPIVVSNNKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 468  
DB 614 AAGVEVTPPIVVSNNKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 661

RESULT 3  
Q88ZV9 PRELIMINARY; PRT; 651 AA.

ID Q88ZV9; PRELIMINARY; PRT; 651 AA.  
AC Q88ZV9; PRELIMINARY; PRT; 651 AA.  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Sucrose PTS, EIIBCA (EC 2.7.1.69).  
GN Name=pts1BCA; OrderedLocusNames=lp\_0185;  
OS Lactobacillus plantarum.  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OX NCBI\_TaxID=1590;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCIMB 8826 / WCFS1;  
RX MEDLINE=22480296; PubMed=12566566;  
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,  
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,  
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,  
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,  
RA De Vos W.M., Siezen R.J.;  
RA "Complete genome sequence of Lactobacillus plantarum WCFS1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).  
DR EMBL; AL935252; CAD62855.1; -.  
DR HSSP; P20166; 1GPR.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0009382; F:protein-N(Pi)-phosphohistidine-sugar phospho. . .; IEA.  
DR GO; GO:0005351; F:sugar porter activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR011055; Dup hybrid motif.  
DR InterPro; IPR001996; Ptrans\_EIIB.  
DR InterPro; IPR003352; Ptrans\_EIIC.

DR TIGR00830; PTB8; 1.  
DR TIGR01996; PTS-II-BC-sucr; 1.  
DR PROSITE; PS01035; PTS\_EIIB\_CYS; 1.  
KW Complete proteome; Transferase.  
SQ SEQUENCE 661 AA; 69148 MW; 69148 MW; COE5E1C7833B9F74 CRC64;

Query Match 100.0%; Score 2363; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 3.6e-144;  
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVFPVLNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 60  
DB 194 MAMVFPVLNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 253  
QY 61 KRLKGTADFLITPVLITLLTGFLTFIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFLGL 120  
DB 254 KRLKGTADFLITPVLITLLTGFLTFIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFLGL 313  
QY 121 VYSPVITGLHOSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 180  
DB 314 VYSPVITGLHOSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 373  
QY 181 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 240  
DB 374 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 433  
QY 241 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEABAP 300  
DB 434 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEABAP 493  
QY 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIIVAFP 360  
DB 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIIVAFP 553  
QY 361 SGHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELICEFDIDA 420  
DB 554 SGHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELICEFDIDA 613  
QY 421 AAGVEVTPPIVVSNNKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 468  
DB 614 AAGVEVTPPIVVSNNKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 661

RESULT 2  
CAF21304 PRELIMINARY; PRT; 661 AA.

ID CAF21304; PRELIMINARY; PRT; 661 AA.  
AC CAF21304; PRELIMINARY; PRT; 661 AA.  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE ENZYME II SUCROSE PROTEIN (EC 2.7.1.69).  
GN PTSS OR CG2925.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriidae; Corynebacterium.  
OX NCBI\_TaxID=1718;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RX PubMed=12948626;  
RA Kalinowski J., Bache B., Bartels D., Bischoff N., Bott M.,  
RA Burkowski A., Busch N., Eggeling L., Birkmanns B.J., Gaigalat L.,  
RA Goemann A., Hartmann M., Huchmacker K., Kraemer R., Linke B.,  
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puchler A.,  
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,  
RA Tauch A.;  
RA "The complete Corynebacterium glutamicum ATCC 13032 genome sequence  
RT and its impact on the production of L-aspartate-derived amino acids  
RT and vitamins.";  
RL J. Biotechnol. 104:5-25(2003).  
DR EMBL; BX927155; CAF21304.1; -.  
KW Transferase.  
SQ SEQUENCE 661 AA; 69148 MW; COE5E1C7833B9F74 CRC64;

DR InterPro; IPR001127; PTS\_EI1A.  
 DR InterPro; IPR010973; PTS\_EI1A-1; 1.  
 DR Pfam; PF00358; PTS\_EI1A-1; 1.  
 DR Pfam; PF00367; PTS\_EI1B; 1.  
 DR Pfam; PF02378; PTS\_EI1C; 1.  
 DR ProDom; PD001476; Ptrans\_EI1B; 1.  
 DR ProDom; PD002243; Ptrans\_EI1A; 1.  
 DR TIGRFAMs; TIGR00830; PTBA; 1.  
 DR TIGRFAMs; TIGR01996; PTS\_EI1A-1; 1.  
 DR PROSITE; PS00371; PTS\_EI1A-1; 1.  
 DR PROSITE; PS01035; PTS\_EI1B\_CYS; 1.  
 KW Complete proteome; Transferrase.  
 SQ SEQUENCE 651 AA; 68514 MW; CBF9535EDDB73DD6 CRC64;

Query Match 45.0%; Score 1062.5; DB 2; Length 651;  
 Best Local Similarity 46.7%; Pred. No. 2.7e-60;  
 Matches 221; Conservative 80; Mismatches 149; Indels 23; Gaps 5;

QY 1 MAMVPSLVNGVDVAATMAAGEMPMSLGLDVAQAGYQGVLPVLVSVWILATIEKFLH 60  
 Db 194 MIMVPSLVNGYSVATTMAAGKQVYVNVFGLHVAQAGYQGVLPVLAVAYILATLEKFFH 253  
 QY 61 KRLKGTADFLITPVLTLLTGLTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG 120  
 Db 254 KHKGAFDFTFPMFAIVITGFTFIVGPVLTVDALNTGLVGLYNTGIMGIFGL 313  
 QY 121 VYSPVITGLHQSPPIELF-----NOGGSFIPATASMANIAQGAACLAIVFFLAKSEKL 176  
 Db 314 LYSIVITGLHQTPAETQLLANVAKTGGSFIPVASMANIQGAATLAIFPAKSKQ 373  
 QY 177 KGLAGSGVSAVLGITEPAIFGVNLRWPFPIGIGTAAIGALIALFNKAVLGAAGF 236  
 Db 374 KALTSSAGVSALLGITTEPAIFGVNLRWPFPIGIGTAAIGALIALFNKAVLGAAGF 236  
 QY 237 LGVVSIDAPDMWFLVCVVTFITAFGAATAYGLYLVNRNGSIDPDATAPVAGTTKAE 296  
 Db 434 LGFISIAKSPAPFMSLAVISFVAIFITFY-----AKTLGDGDDQVKSAPTSTV--- 486  
 QY 297 AEAPAEFNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKGOLVSPVSGKIV 356  
 Db 487 -----INVNDIISAPVTGTSLSKLVNDQVFSAEIMGKGAIVPSADQVAPADGVIT 540  
 QY 357 VAFPSGHAPVATKADGSDNVILMHIGFDVTNGLTHNPPLKGGDEVKAGELICEPDI 416  
 Db 541 VTYDSHHAIGKNTA-----GABILHLGLDVTNGLNGEFTTNVQKGDVTHQGDLLGTEVD 596  
 QY 417 DAIKAAGYEVTTPIVSVNKKTPVNTYGLGIEAGANLLNVAKKEA--VPAT 467  
 Db 597 AALKAANDPTVMTLVNTTANYANVERLKVTVNQAGEQLVALTEPTASSVAAT 649

## RESULT 4

PTSA PEDPE  
 ID PTSA PEDPE STANDARD; PRT; 651 AA.  
 AC P43470;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE PTS system, sucrose-specific IIABC component (BIIABC-Scr) (Sucrose-  
 DE permease IIABC component) (Phosphotransferase enzyme II, ABC  
 DE component) (EC 2.7.1.69) (BII-Scr).  
 GN Name=sra;  
 OS Pedicoccus pentosaceus.  
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pedicoccus.  
 OX NCBI\_TaxID=1255;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=PP1.0;  
 RC Leenhouts K.K.J., Bolhuis A.A., Kok J.J., Venema G.G.;  
 RA "The sucrose and raffinose operons of *Pedicoccus pentosaceus*  
 RT PP1.0";  
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: This is a component of the phosphoenolpyruvate-dependent

sugar phosphotransferase system (PTS), a major carbohydrate active  
 -transport system. The IICD domains contain the sugar binding site  
 and the transmembrane channel; the IIA domain contains the primary  
 phosphorylation site (the donor is phospho-HPr); IIA transfers its  
 phosphoryl group to the IIB domain which finally transfers it to  
 the sugar.  
 -!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein  
 histidine + sugar phosphate.  
 -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 -!- SIMILARITY: Contains 1 PTS\_EI1A domain.  
 -!- SIMILARITY: Contains 1 PTS\_EI1B domain.  
 -!- SIMILARITY: Contains 1 PTS\_EI1C domain.  
 -!- SIMILARITY: Contains 1 PTS\_EI1C domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Z32771; CAA83668.1; -.  
 EMBL; U32093; AAA25567.1; -.  
 PIR; S44257; S44257.  
 HSP; P20166; IGPR.  
 InterPro; IPR011055; Dup hybrid motif.  
 InterPro; IPR001996; Ptrans\_EI1B.  
 InterPro; IPR003352; Ptrans\_EI1C.  
 InterPro; IPR001127; PTS\_EI1A.  
 InterPro; IPR010973; PTS\_EI1A-1; 1.  
 Pfam; PF00358; PTS\_EI1A-1; 1.  
 Pfam; PF00367; PTS\_EI1B; 1.  
 Pfam; PF02378; PTS\_EI1C; 1.  
 ProDom; PD002243; PTS\_EI1A; 1.  
 ProDom; PD001476; PTS\_EI1B; 1.  
 TIGRFAMs; TIGR00826; EI1B\_glc; 1.  
 TIGRFAMs; TIGR00830; PTBA; 1.  
 TIGRFAMs; TIGR01995; PTS\_EI1A-1; 1.  
 TIGRFAMs; TIGR01996; PTS\_EI1B-1; 1.  
 TIGRFAMs; TIGR01996; PTS\_EI1C-1; 1.  
 PROSITE; PS00371; PTS\_EI1A-1; 1.  
 PROSITE; PS01035; PTS\_EI1B\_CYS; 1.  
 KW Phosphorylation; Phosphotransferase system; Sugar transport;  
 Transferrase; Transmembrane.  
 FT DOMAIN 1 40 EI1B.  
 FT DOMAIN 111 418 EI1C.  
 FT DOMAIN 510 651 EI1A.  
 FT MOD RES 25 25 Phosphocysteine (By similarity).  
 FT MOD RES 324 324 Phosphohistidine (By similarity).  
 FT MOD RES 562 562 Phosphohistidine (By similarity).  
 SQ SEQUENCE 651 AA; 68454 MW; C87BA09D550A77F8 CRC64;

Query Match 44.8%; Score 1058.5; DB 1; Length 651;  
 Best Local Similarity 46.7%; Pred. No. 4.9e-60;  
 Matches 221; Conservative 78; Mismatches 151; Indels 23; Gaps 5;

QY 1 MAMVPSLVNGVDVAATMAAGEMPMSLGLDVAQAGYQGVLPVLVSVWILATIEKFLH 60  
 Db 194 MIMVPSLVNGYSVATTMAAGKQVYVNVFGLHVAQAGYQGVLPVLAVAYILATLEKFFH 253  
 QY 61 KRLKGTADFLITPVLTLLTGLTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG 120  
 Db 254 KHKGAFDFTFPMFAIVITGFTFIVGPVLTVDALNTGLVGLYNTGIMGIFGL 313  
 QY 121 VYSPVITGLHQSPPIELF-----NOGGSFIPATASMANIAQGAACLAIVFFLAKSEKL 176  
 Db 314 LYSIVITGLHQTPAETQLLANVAKTGGSFIPVASMANIQGAATLAIFPAKSKQ 373  
 QY 177 KGLAGSGVSAVLGITEPAIFGVNLRWPFPIGIGTAAIGALIALFNKAVLGAAGF 236  
 Db 374 KALTSSAGVSALLGITTEPAIFGVNLRWPFPIGIGTAAIGALIALFNKAVLGAAGF 236  
 QY 237 LGVVSIDAPDMWFLVCVVTFITAFGAATAYGLYLVNRNGSIDPDATAPVAGTTKAE 296

Db 434 IGISTASKSIPAFMLSAVISFVVAFLPTIY-----AKRTLGDDRDVOVKSPATISV--- 486

QY 297 AEAPAFSNDSTIIQAPLTGEALIALSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIV 356

Db 487 -----INVNDIISAPVTGASESLKQVNDQVFSAEIMGKGAALVPSDDQVAPADGVIT 540

QY 357 VAPFSGHAFVRKAEKDSNVDILMHIGPTVMINGTHFNPLKKQDGEVKAAGELLCEFDI 416

Db 541 VTYDSHAYGIKTTA-----GAEILIHGLDVTNNGEHHFTTNVQKGDVTHQGDLLGTFDI 596

QY 417 DAUKAAGVEVTPVIVGSKYKTPVNTYGLGEIEAGANL--LNVAKKEAVPAT 467

Db 597 AALKAANYDPTVMLIVTINTANYANVERLKVNTVQAGEQLVALTAPAASSVAAT 649

RESULT 5

Q93ML1 ID Q93ML1 PRELIMINARY; PRT; 534 AA.

AC Q93ML1; 2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Sucrose-specific enzyme II of the PTS (Fragment).

GN Name=sucrA;

OS Lactobacillus sakei.

OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;

OC Lactobacillus.

OX NCBI\_TaxID=1599;

EN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21822499; PubMed=11832506;

RA Duzek A.-M., Chailou S., Zagorec M.;

RA Duzek A.M., Chailou S., Hissler L., Stentz R.,

RA Champoinier-Verges M.C., Albert C.A., Zagorec M.;

RT "Physical and Genetic map of the Lactobacillus sakei 23K chromosome.";

RL Microbiology 148:421-431(2002).

RN [2]

RP SEQUENCE FROM N.A.

RX Duzek A.-M., Chailou S., Zagorec M.;

RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF401046; AAK92528.1; -.

DR HSSP; P45618; 2GPR.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008982; F:protein-N(P1)-phosphohistidine-sugar phospho. . .; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR011055; Dup hybrid motif.

DR InterPro; IPR003352; Ptrans EIIC.

DR InterPro; IPR001127; PTS EIIA.

DR Pfam; PF00358; PTS\_EIIA\_1; 1.

DR Pfam; PF02378; PTS\_EIIC; 1.

DR TIGRFAMs; TIGR00830; PTBA; 1.

DR PROSITE; PS00371; PTS\_EIIA\_1; 1.

FT NON\_TER

SQ SEQUENCE 534 AA; 56230 MW; 01804F9DE70C0089 CRC64;

Query Match 41.3%; Score 977; DB 2; Length 534;

Best Local Similarity 44.2%; Pred. No. 7.3e-55;

Matches 211; Conservative 90; Mismatches 138; Indels 38; Gaps 10;

QY 1 MAWVPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTVLPLVLSWILATIEKFLH 60

Db 81 MMLVMPNLVNGYGVASISATGHTYWHVFLNIAQAGYQGTVPVIGVAFILANLEKFFH 140

QY 61 KELKGFADFLITPLVLTLLTGLTFTIAIGFAMRWGVDVLHAGLQGHYDFGPGVGLLFL 120

Db 141 KHLNDAVDFTFPMLSIITGLTFTTLVGLPALIVSGVTDLSVWAYQTLGAVGMGIFGL 200

QY 121 VYSPVITGLHQSFPIELF----NQGSFTFATASMANIAQGAACLAIVFLAKSEKL 176

Db 201 GYSAIVLTGLHQSFPAIETLLADIARTKGTGSFFFPVAAMANTIAQGAATFAVFEVTKNQ 260

QY 177 KGLAGSGVSAVLGITETPAIFGVNLRWRPFFTGICTPAATGALLIALFNKAVAGAGF 236

Db 261 KSLTTSAGISAMLGITETPAIFGVNLRKLPDFFTGILIASGSSFIIGLLHLSVSMGPAGI 320

QY 237 LGVVSIDAPDMVNFVCAVVTFFIAGAAATAYGLYLVRNGSIDPDATAPVPAAGTTKAE 296

Db 321 IGFIATAPKSIPSPFMGALISFVIAFYGTLYGKKAMK-----TTTEE 363

QY 297 A--EAP-----EFSNDSTIIQAPLTGEALIALSVSDAMFASGKLGSGVAIVPTKGOLVS 349

Db 364 IINEAPATPEVRLQDEK--ISAPVTGRIVDLASVPDPVASEAMGKIALIMPSTQDVLA 422

QY 350 PVSGKIIVAFPSGHAFVRKAEKDSNVDILMHIGPTVMINGTHFNPLKKQDGEVKA 409

Db 423 PVTGVITIAANTGHAYGI--KSDDGA--EVLIIHLGLDVTNNGIGTGFKEIVQGGQHVSEGD 478

QY 410 LLCEFDIDAIAAGYEVTTPIVSN---YKKTGPVNTYGLGEIEAGANLNV-AKKE 462

Db 479 LLGHFDIDKIKQAGLTFTMTIVNTAGYAQVDPLLTVDKAAHQ--GEEIIQLHAKKD 534

RESULT 6

Q9S6S6 ID Q9S6S6 PRELIMINARY; PRT; 650 AA.

AC Q9S6S6;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Enzyme II sucrose protein (EC 2.7.1.69).

GN Name=sacB;

OS Lactococcus lactis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI\_TaxID=1358;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99173919; PubMed=10074089;

RA Luesink E.J., Marugg J.D., Kuipers O.P., De Vos W.M.;

RT "Characterization of the divergent sacBK and sacAR operons, involved in sucrose utilization by Lactococcus lactis.";

RL J. Bacteriol. 181:1924-1926(1999).

DR EMBL; Z97015; CAB09690.1; -.

DR HSSP; P45618; 2GPR.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008982; F:protein-N(P1)-phosphohistidine-sugar phospho. . .; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0016740; P:transferase activity; IEA.

DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR011055; Dup hybrid motif.

DR InterPro; IPR001996; Ptrans EIIB.

DR InterPro; IPR003352; Ptrans EIIC.

DR InterPro; IPR001127; PTS EIIA.

DR InterPro; IPR010373; PTS\_EIIA\_1; 1.

DR Pfam; PF00358; PTS\_EIIA\_1; 1.

DR Pfam; PF02378; PTS\_EIIC; 1.

DR Pfam; PF00367; PTS\_EIIB; 1.

DR ProbDom; PD001476; Ptrans EIIB; 1.

DR ProbDom; PD002243; PTS\_EIIA; 1.

DR TIGRFAMs; TIGR00830; PTBA; 1.

DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.

DR PROSITE; PS00371; PTS\_EIIA\_1; 1.

DR PROSITE; PS01035; PTS\_EIIB\_CYS; 1.

KW Transferase.

SQ SEQUENCE 650 AA; 69636 MW; 250E26F7664D4204 CRC64;

Query Match 40.7%; Score 962; DB 2; Length 650;

Best Local Similarity 42.9%; Pred. No. 8.2e-54;

Matches 203; Conservative 89; Mismatches 145; Indels 36; Gaps 7;

QY 1 MAWVPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTVLPLVLSWILATIEKFLH 60

Db 194 MLNVPELVNGVNAEAIASHNHTYWDIFGKVAQAGYQGTVPVIGVAFILAKRFFH 253



```
QY 228 AVALGAGFGLGVSVISIDAPDMVMFVCAVVTFFIAFGAIAAYGLYLVRNGSIDPDATAAP 287
Db 434 AVSLGAGFGLGFSIKASSIPFYWCCELISFAIAFAVITYGYG-----KTKAVDVFAEAA 488
QY 288 VPAGTTKABEAAPAEFNSDST-----ITQAPLTGEAIALSSVSDFAMFASGKLGSVAI 340
Db 489 VERAIEVQ-EIPEEAASAANKAQVTEVLAAPLAGEAVELTSVNDPVFSSEANGKGIAI 547
QY 341 VPTKGQLVSPVSGKIVVAFPSGHAFVTRKADGNSVDILMHIGFTDVNLNGTHFNPLK 400
Db 548 KPSGNTVYAPVDGTQVIAFTGHAYGI--KSDNGA--EILIHIGIDTVSMEGKGFQKVQ 603
QY 401 QGDEVKAGELLCEPDIDAIKAAGYEVTTPVVSNNYKKTGPNTY-GLGEIEAGANLINVA 459
Db 604 ADQKIKKGDVLGTFDSKIAEAGLDNTTWFIVTNTADYASVETLIASSGTVAVGDSLLEVK 663
QY 460 K 460
Db 664 K 664

RESULT 8
Q74HI8
ID Q74HI8 PRELIMINARY; PRT; 647 AA.
AC Q74HI8;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE Phosphoenolpyruvate-dependent sugar phosphotransferase system EIIABC,
DE sucrose specific.
GN OrderedLocusNames=LJ0519;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017206; AAS09702.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR011055; Dup hybrid motif.
DR InterPro; IPR011535; EIIH_glc.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR003352; Ptrans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR010973; PTS_II_BC_sucr.
DR Pfam; PF00358; PTS_EIIA.1; 1.
DR Pfam; PF00367; PTS_EIIB.1.
DR Pfam; PF02378; PTS_EIIC.1.
DR ProDom; PD001476; Ptrans_EIIB.1.
DR ProDom; PD002243; PTS_EIIA.1.
DR TIGRFAMs; TIGR00826; EIIH_glc.1.
DR TIGRFAMs; TIGR00830; PTBA.1.
DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr.1.
DR PROSITE; PS00371; PTS_EIIA.1.1.
DR PROSITE; PS01035; PTS_EIIB_CUS.1.
DR Complete proteome; Pyruvate; Transferrase.
SQ SEQUENCE 647 AA; 69202 MW; 95FA74C533CE9541 CRC64;

Query Match 38.7%; Score 914.5; DB 2; Length 647;
Best Local Similarity 41.3%; Pred. No. 9.5e-51;
Matches 195; Conservative 88; Mismatches 160; Indels 29; Gaps 6;

QY 1 MANVFPSSLVNGYDVAAATMAAGEMPWMSLFGDLVAQAGYQGTVPVLVSVWILATIEKFLH 60
Db 194 MIMTTPAL-----GGAGKYWDIFGLHVSQTNVQYQVIVPLVAVWLLAFLEKRFH 242
QY 61 KRLKGTADFLITPVLTLTLTGLTFTFAIGPAMRWGVDVLAHQGLQGLYDFGPGVGLLFL 120
Db 243 KWLPSAVDFTFTLLSIMITGLTFTIIGVPFKGVSDAITAAIWLNTTGAFGMGIFGL 302
QY 121 VTSPIVITGLHOSFPPIELELFNQ-----GGSFIFATASMANIAQGAACLAFFFLAKSE 174
```

```
Db 194 MIMTTPAL-----GGAGKYWDIFGLHVSQTNVQYQVIVPLVAVWLLAFLEKRFH 242
QY 61 KRLKGTADFLITPVLTLTLTGLTFTFAIGPAMRWGVDVLAHQGLQGLYDFGPGVGLLFL 120
Db 243 KWLPSAVDFTFTLLSIMITGLTFTIIGVPFKGVSDAITAAIWLNTTGAFGMGIFGL 302
QY 121 VTSPIVITGLHOSFPPIELELFNQ-----GGSFIFATASMANIAQGAACLAFFFLAKSE 174
Db 303 SYSAIVTTGLHQSPFAVETQLLAEYARGRSGDFIFVTACMANVQAQAAFTFAIYFLTKNE 362
QY 175 KLKGLGASGVAVLGITTEPAIFGNLRLRPPFPGITAAIGGALIALFNKIKAVALLCAA 234
Db 363 KVKGLASSGVSALLGITTEPALFGVNLKYKFPFPCALIGVGAAPAGLHMVHTAALGSA 422
QY 235 GFLGVSIDAPDMVMFVCAVVTFFIAFGAIAAYGLY-----LVVRNGSIDPDATAAPVP 289
Db 423 GFLGFLSMVFKSIPMWALSIVSIFVAFGLTFVYGRHKFKEDVVEESGTEV---SAGDQV 479
QY 290 AGTTKAEAPAPAFBSNDSTIIQAPLTGEAIALSSVSDFAMFASGKLGSVAIVPTKGQVLS 349
Db 480 AQQEKARQIIKEDKELHDEVIAAPVSGKABSLKDNDVDFSTFSTANGKGAANVPSEGTIYS 539
QY 350 PVSGKIVVAFPSGHAFVTRKADGNSVDILMHIGFTDVNLNGTHFNPLKQSGDEVKAGE 409
Db 540 PVTGEITVAVETKHAYGI--KSDEGA--EVLHIGIDTVNMKGSGFTDVKQGGHVEKGE 595
QY 410 LLCEFDIDAIIKAAGYEVTTPVVSNNYKKTGPNTYGLGREAGANLINVAKK 461
Db 596 KLGTVDDLDAVKKAGYDITVMTVIINTTSYANVQRIDGVEKKHGDLLIAVTKR 647

RESULT 9
AAS09702
ID AAS09702 PRELIMINARY; PRT; 647 AA.
AC AAS09702;
DT 02-MAR-2004 (TremBLrel. 27, Created)
DT 02-MAR-2004 (TremBLrel. 27, Last sequence update)
DE Phosphoenolpyruvate-dependent sugar phosphotransferase system EIIABC,
DE sucrose specific.
GN LJ0519
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017206; AAS09702.1; -.
DR Pyruvate; Transferrase.
SQ SEQUENCE 647 AA; 69202 MW; 95FA74C533CE9541 CRC64;

Query Match 38.7%; Score 914.5; DB 2; Length 647;
Best Local Similarity 41.3%; Pred. No. 9.5e-51;
Matches 195; Conservative 88; Mismatches 160; Indels 29; Gaps 6;

QY 1 MANVFPSSLVNGYDVAAATMAAGEMPWMSLFGDLVAQAGYQGTVPVLVSVWILATIEKFLH 60
Db 194 MIMTTPAL-----GGAGKYWDIFGLHVSQTNVQYQVIVPLVAVWLLAFLEKRFH 242
QY 61 KRLKGTADFLITPVLTLTLTGLTFTFAIGPAMRWGVDVLAHQGLQGLYDFGPGVGLLFL 120
Db 243 KWLPSAVDFTFTLLSIMITGLTFTIIGVPFKGVSDAITAAIWLNTTGAFGMGIFGL 302
QY 121 VTSPIVITGLHOSFPPIELELFNQ-----GGSFIFATASMANIAQGAACLAFFFLAKSE 174
```

```
Db 303 SYSAIVTTGLHQSFPFAVETOLLAEYAGRGSGDFIFVTACMANVAQGAATFAIYFLTKNE 362
QY 175 KIKGLAGASGVSAVLGITEPAIGFVNLRLRPPFFIGTAAIGCALIALFNKKAVALGAA 234
Db 363 KVKGLASSGVSAVLLGITEPALGVNLKYKFFPFCALIGAGVAAAGAGLHVTAALGSA 422
QY 235 GFLGVVSDADPMWMLVCAVVTFFIAFGAAIAYGLY-----LVRNGSIDDPDATAAPVP 289
Db 423 GFLGFLSMVFKSPIMWALSIVISFIAFGLTVYKGRHFKREDVVERSGTVE---SAGDQV 479
QY 290 AGTTKAEABAPAEFSDNSTIIQAPLTGEATASVSDAMFASGKLSGVVAIVTKGOLVS 349
Db 480 AQEKAEQIIEKDELHDEVAIPVSKAESLKDNDVPFESTAMGGAAMVSEGTIYS 539
QY 350 PVSQKIVVAPPSGHAFVAVRTKABDGSNDVILMHIGFDTVNLNGTHFNPLKKGQDEVKAGE 409
Db 540 PVTGEITVAETKHAYGI--KSDEGA--EVLHIGLDTVNMKGEGFTYDKQQQHWKEGE 595
QY 410 LLCEFDIDAIAKAGYEVTTPVIVSNYKKGTPVNTYGLGEIEAGANLLNVAKK 461
Db 596 KLGTVDLDAVKKAGYDTTWTMTVTNTTSYANVQRIDGVEKKHGDLLIAVTKR 647

RESULT 10
Q7WWP7 PRELIMINARY; PRT; 649 AA.
AC Q7WWP7;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Sucrose PTS transporter (EC 2.7.1.69).
GN Name=scrA;
OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]_TaxID=1579;
RP SEQUENCE FROM N.A.
RC STRAIN=NCFM;
RX PubMed=12847288;
RA Barrangou R., Altermann E., Hutkins R., Cano R., Klaenhammer T.R.;
RT "Functional and comparative genomic analyses of an operon involved in
fructooligosaccharide utilization by Lactobacillus acidophilus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8957-8962(2003).
DR EMBL; AY177419; AAC38866.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(Pi)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001996; P:trans_EIIB.
DR InterPro; IPR003352; P:trans_EIIC.
DR InterPro; IPR01127; PTS_EIIA.
DR InterPro; IPR010973; PTS_II_BC_sucr.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; P:trans_EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; P:trans_EIIB; 1.
DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
DR Transferrase.
SQ SEQUENCE 649 AA; 59446 MW; B4377CA24D00507B CRC64;

Query Match 38.1%; Score 901; DB 2; Length 649;
Best Local Similarity 40.8%; Pred. No. 7.1e-50;
Matches 183; Conservative 81; Mismatches 149; Indels 36; Gaps 6;

QY 1 MAMVFPSPVINGVDVAATWAAGEMPWWSLFGLDVAQAGQGVTLVPLVSVWILATIEKPLH 60
```

```
Db 194 MIMTTPGL-----CGATKFWDFI FGLHVAQTNQYQVTPVLVAVVWLSIFEKVFH 242
QY 61 KRUKGTADFLITVLTLLTGLFTTFAIGPAMRWGVDLAHLGQLGYDFGPGPVGLLFLGL 120
Db 243 KHLPSAVDFTFITPLLSIMITGFTTFTIGPVFKGVSDAITNAIWLVDYTTGAFGCVFGL 302
QY 121 VYSPVITGLHQSFPPIELELF-----NQGSFIFATASMANIAGAACLAFFLAKS 173
Db 303 SYSAIVTTGLHQSFPFAVETOLLAEAFKPNPASSGDFIFVTACMANVAQGAATFAIYFLTKN 362
QY 174 EKLKGLAGASGVSAVLGITEPAIGFVNLRLRPPFFIGTAAIGCALIALFNKKAVALGAA 233
Db 363 KMKGLASSGVSAVLLGITEPALGVNLKYKFFPFCALIGSGVAAAFAGLHVTAALGSA 422
QY 234 AGPLGVVSDADPMWMLVCAVVTFFIAFGAAIAYGLYLVRNGSIDDPDATAAPVPAG-- 291
Db 423 AGFLGFLSYPTIPMPVSVSAISFAVSFLITVYVG-----KSHLKEEVAEQDVPVDA 476
QY 292 -----TTKAEABAPAEFSDNSTIIQAPLTGEATASVSDAMFASGKLSGVVAIVTKG 345
Db 477 TDYAEITQAEKIGKEQLAKDBEIIYSPVDGTPESTLKVNDQVFSAKLMGDGAIVPSDG 536
QY 346 OLVPVSGKIVVAPPSGHAFVAVRTKABDGSNDVILMHIGFDTVNLNGTHFNPLKKGQDEV 405
Db 537 TIVAPVTGVTIAYETKHAYGI--KSDDGA--EVLHIGLDTVNLKGEHPESVFKGQQRV 592
QY 406 KAGELLCEFDIDAIAKAGYEVTTPVIVSN 434
Db 593 EKGDKLSGVYDLDAVKKAGYDTTVMVVITN 621

RESULT 11
Q82YR5 PRELIMINARY; PRT; 643 AA.
AC Q82YR5;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE PTS system, IIABC components.
GN OrderedLocNames=EFA0067;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]_TaxID=1351;
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
Enterococcus faecalis.";
RL Science 299:2071-2074(2003).
DR EMBL; AB016833; AAC83060.1; -
DR HSSP; P20166; IGPR.
DR TIGR; BFA0067; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(Pi)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR01055; Dup hybrid motif.
DR InterPro; IPR001996; P:trans_EIIB.
DR InterPro; IPR003352; P:trans_EIIC.
DR InterPro; IPR01127; PTS_EIIA.
DR InterPro; IPR010973; PTS_II_BC_sucr.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
```

DR	Pfam; PF02378; PTS_EIIC; 1.	
DR	ProDom; PD001476; Ptrans_EIIB; 1.	
DR	ProDom; PD002243; PTS_EIIA; 1.	
DR	TIGRFAMS; TIGR00830; PTBA; 1.	
DR	TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.	
DR	PROSITE; PS00371; PTS_EIIA.1; 1.	
DR	PROSITE; PS01035; PTS_EIIB_Cys; 1.	
KW	Complete proteome; Plasmid.	
SEQ	SEQUENCE 643 AA; 68532 MW; 752F65D8154AA06C CRC64;	

  

Query Match	35.2%; Score 832.5; DB 2; Length 643;
Best Local Similarity	39.3%; Pred. No. 1.8e-45;
Matches 184; Conservative 84; Mismatches 171; Indels 29; Gaps 7;	

  

Qy	1	MAMVFPSSLVNGYDYAATWAAGEMPWSLFGLDVAQAGYQGQGVLPVLPVIVVSWLATIEKFLH 60
Db	194	MIMTTPDL-----GGATYWNIFGVHVVAQTNAYOVIPVLASVYLSLEIKYFH 242
Qy	61	KRLKGTADFLLTPVILLLLGLFTFIAIGPAMRWGDVLAHGLGCLGYDFGGFVGGLLFLGL 120
Db	243	KKLPSIDFTFTPLLSVIITGLFTFTVIGPTMLLSNGITDAIVWLYNATFGIGMGIFGG 302
Qy	121	VYSPIVTGLHQSPPIELELFNQ-----GGSFATASMANIAOAGACLAVFLAKSE 174
Db	303	TYSLIVMTGLHQSPFAIETQLLSAWTNGIGHGDFIFVVASMANVAQGAATFAIWFLTKNS 362
Qy	175	KLKGAGAGSVSAVLGITTEPAIFGNLRLRPPFFIGITGAAGGALIALFNKIVALGAA 234
Db	363	KTKLSSAGSALSLGITTEPAIFGNLRLRPPFFCALIGSGIAAATGLKKVAISLGA 422
Qy	235	GFLGVSDIDAPDMWNFLVCVAVTFPIAFGAATYGLYLVRNGSIDDPATAAPVPAGTTK 294
Db	423	GFLGFLSINATSIPIYLCELLSFSVTALTIFYG--RTRSSIFAAEAIAEQTSDVTSE 480
Qy	295	AEAEAPAFSNDSTI--TOAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKGQLVSPVS 352
Db	481	INTNQIANADEPTTIVETIVSPLAGETITLGSVNDPVSFSESIGKGAIKPNGNTIYSPVD 540
Qy	353	GKIVVAFPSGHAFATVRAEDGSNDVILMHGFDVNLNGTHFNLPKKQG--DEVKAGEL 410
Db	541	GIVQVVFETGHAYDLKS---NTGAELIHLHGIDIVSLNGKFT--KKVGAQKQVKKEGV 594
Qy	411	ICEFDIDAKAAGYEVVTPVIVSNVKKTGPNVNTYGLGTEIAGANLINV 458
Db	595	LGTFFSTVITNSGLDDTMMVIVTNSKDYSEVIPKTIIVTEGAALITI 642

  

RESULT 12	
Q9L8G6	PRELIMINARY; PRT; 627 AA.
AC	Q9L8G6; PRELIMINARY;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Scra (Fusion: PTS system, beta-glucosides specific IIABR component).
GN	Name=scra; OrderedlocusNames=CAC0423;
OS	Clostridium acetobutylicum.
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC	Clostridium.
OX	NCBI_TaxID=1488;
RN	[1]_TaxID=1488;
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 824;
RX	MEDLINE=20391269; PubMed=10937490;
RA	Tangney M., Mitchell W.J.;
RT	"Analysis of a catabolic operon for sucrose transport and metabolism
RT	in Clostridium acetobutylicum ATCC 824.";
RL	J. Mol. Microbiol. Biotechnol. 2:71-80(2000).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX	MEDLINE=21359325; PubMed=11466286;

RESULT 13

DOI=10.1128/JB.183.16.4823-4838.2001;  
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.J., Sabathe F., Douchette-Stamm L.A., Soucaille P.,  
 RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.,  
 RA "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium *Clostridium acetobutylicum*,"  
 RL J. Bacteriol. 183:4823-4838 (2001).  
 DR EMBL; AF205034; AAF35839.1; -;  
 DR EMBL; AF205034; AAF35839.1; -;  
 DR FIR; H96951; AAK78403.1; -;  
 DR HSP; P20166; IGPR.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0008982; F:protein-N(P)-phosphochistidine-sugar phosph. .; IEA.  
 DR GO; GO:0003351; F:sugar porter activity; IEA.  
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.  
 DR GO; GO:0006810; P:transpor; IEA.  
 DR InterPro; IPR011055; Dup hybrid motif.  
 DR InterPro; IPR011254; Peptidase S1.  
 DR InterPro; IPR001996; P:trans EIIB.  
 DR InterPro; IPR003352; P:trans EIIC.  
 DR InterPro; IPR001127; P:trans EIIA.  
 DR InterPro; IPR010973; P:trans EII BC\_sucr.  
 DR Pfam; PF003358; P:trans EIIA\_1; I.  
 DR Pfam; PF00367; P:trans EIIIB; I.  
 DR Pfam; PF02378; P:trans EIIIC; I.  
 DR ProDom; PD001476; P:trans EIIB; I.  
 DR ProDom; PD002243; P:trans EIIA; I.  
 DR TIGRFAMS; TIGR00830; P:trans EIIA; I.  
 DR TIGRFAMS; TIGR001996; P:trans EII BC\_sucr; I.  
 DR PROSITE; PS00371; P:trans EIIA\_1; I.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 627 AA; 67596 MW; B835AB0238FCA436 CRC64;  
 Query Match 32.1%; Score 759.5; DB 2; Length 627;  
 Best Local Similarity 36.7%; Pred. No. 9.2e-41;  
 Matches 163; Conservative 80; Mismatches 176; Indels 25; Gaps 5  
 QY 3 MVFSLVNGYDVAATMAAGEMPMSLFGDLVQAQGYQGVTVPLVLSVILATIEKFLHKR 62  
 DB 186 MIHPDLQNAW---TLGEGIKHTINIFGLNIGMVGYQGVTVPLILLSVMSVIEKGLRK 241  
 QY 63 LKGTADELITPVLTLLLTGFTFTTAIGPAMWVGDLVAHGLQGLVDFGSPVGLGLFLGLV 122  
 DB 242 VPEALDILLTFPLMTITGTFAMVVGPGFVGDEISGLIQLTLINTTGFSGVLFGGLY 301  
 QY 123 SPIVITGLHQSPFPIELELEFNQGG---SFIFATASMANIAQGAACIAYFFLAKSEKLKGL 179  
 DB 302 SLIVITGHHFSFAIEAGLLANPAHKNFLLPIWSMANVQAQGAALAVFKTRDKWKSI 361  
 QY 180 AGASGVSAVLGITPEIPAIGVNLRLRWPFGIIGTHTAIGALIALFNKIKALCAAGFLGV 239  
 DB 362 AAPASFCLLGITPEIPAIGVNLRLYTKPIAGALGAIGGGYIVFTKVMAMTAGVGTIGPI 421  
 QY 240 VSIDAPDMVMTLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTITKAEABA 299  
 DB 422 AIVKQGSFLINTIAMI-----IAPGAGFIAMVLGIKEITTEEDLNKETVYNDIKVEEVES 477  
 QY 300 PAEFSNDSTIIQAPLTGGBAIALSSVSDAMFASGKIGSGVAIVPTTGQLVSPVSGKIIVAF 359  
 DB 478 -----VVSPVNGKVLIIKNVPDKTFAEGLIGDGIQVDPEDGEVVSPIDGTVVHV 527  
 QY 360 PSGHAFAYRTKAEDSGNDVILMHGIFDTVNLNGTHFNPLKKGDEYKAGCELCEFDIAI 419  
 DB 528 ETKHAIAWRKS-----NGVEMLLHIGIDIVKMEGNGFKSFINDGEVKKGDKLIQFDLDLV 583  
 QY 420 KAAGYEVTTPVIVSNYKKTGPVNT 443  
 DB 584 KEKAVSPVLTIVTNHEDMGFVNS 607

RESULT 13

Q831B4  
 ID Q831B4 PRELIMINARY; PRT; 626 AA.  
 AC Q831B4;  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE PTS system, beta-glucoside-specific IIBC component.  
 GN OrderedLocusNames=EF2598;  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 OX NCBI\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=V583 / ATCC 700802;  
 RX MEDLINE=22550857; PubMed12663927; DOI=10.1126/science.1080613;  
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
 RA Tettelin H., Dodson R.J., Unayun L.A., Brinkac L.M., Beanan M.J.,  
 RA Dougherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,  
 RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,  
 RA Fraser C.M.;  
 RT "Role of mobile DNA in the evolution of vancomycin-resistant  
 RT Enterococcus faecalis";  
 RL Science 299:2071-2074 (2003).  
 DR EMBL; AF016955; AA082308.1; -.  
 DR HSSP; P08837; IGGR.  
 DR TIGR; EF2598; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.  
 DR GO; GO:0005351; F:sugar porter activity; IEA.  
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.  
 DR GO; GO:0008810; P:transport; IEA.  
 DR InterPro; IPR011055; Dup hybrid motif.  
 DR InterPro; IPR001996; Ptrans\_EIIB.  
 DR InterPro; IPR003352; Ptrans\_EIIC.  
 DR InterPro; IPR001127; PTS\_EI1A.  
 DR InterPro; IPR011297; PTS\_II-ABC\_beta.  
 DR Pfam; PF00358; PTS\_EI1A\_1; 1.  
 DR Pfam; PF00367; PTS\_EIIB; 1.  
 DR Pfam; PF02378; PTS\_EIIC; 1.  
 DR ProDom; PD001476; Ptrans\_EIIB; 1.  
 DR ProDom; PD002243; PTS\_EI1A; 1.  
 DR TIGRFAMs; TIGR00830; PTEA; 1.  
 DR TIGRFAMs; TIGR01995; PTS-II-ABC\_beta; 1.  
 DR PROSITE; PS00371; PTS\_EI1A\_1; UNKNOWN\_1.  
 DR PROSITE; PS01035; PTS\_EIIB\_CYS; 1.  
 DR Complete proteome.  
 SQ SEQUENCE 626 AA; 66244 MW; 0E0BF2C4220570B CRC64;

Query Match 30.3%; Score 715; DB 2; Length 626;  
 Best Local Similarity 37.4%; Pred. No. 6.8e-38;  
 Matches 166; Conservative 72; Mismatches 172; Indels 34; Gaps 7;  
 QY 1 MAMVPSLVNGYDVAAATWAGMPNWSLFGLDVAGAGTGVLPVLVSVWILATIEKPLH 60  
 DB 181 MALVHPSI-----TALAGKTSFAGLPVIGSGYTSVLPILAVFAQSYVERFFK 232  
 QY 61 KRLGKATDFLITPVLTLTLTGFTATGTPAMRWGVDVLHAGLQGLYDFGPGVGLLFL 120  
 DB 233 KVIPSLQICVPLAVFLMAPVFLAIGTGTGVDLWLGQYNAIYAFSPIIAGLLMGS 292  
 QY 121 VYSIVITGLHQSPDIELEFNQSGSPIFATASMANIAQGAACIAVFFLAKSEKLGIA 180  
 DB 293 LQWLVMEGMEHGFVPIIMNLITQGGDTWVPMLLPAVIAQGAALAVFLTKVNLKGLA 352  
 QY 181 GASGVSAVLGTEPAFGVNLRLRPPFPGTGAIGGALIALFNKIKAVAGAA----- 234  
 DB 353 LSSSITTFITGTEPTVYGVTLPLKPFIAACIGGIGGAFVAMNHVKNFTFGLVMSLSL 412  
 QY 235 GFLGVSVSDAPDMWMLCAVVTFFIAPGAALVGLVLRNGSIDDPATAPVAGTTK 294  
 DB 413 GFIPAEYKDTAPMTGAIGAGIAFIIFVLT-----FVIRFEDQNPB-----TATEK 460

QY 295 AEAE---APAEFS-NDSTIIQAPLTGEATIALSSVDAMFASGKLGSGVAIVPTKGQLVSP 350  
 DB 461 TETDKMVAIPVKTNQEDKILIASLPQELFLEKVDQPPVFASGALGKGVAIEPTGKLYAP 520  
 QY 351 VSGKIVVAFPSGHAFVRKKAEDGSNVVDILMHITGFTVNLNGTHENPLKKGQDEVKAGEL 410  
 DB 521 ADGEITLTPFTGHAVGLTIT---EGVELLMHLMGMDTVELDGKGFELSVKQGSVKKGGDL 576  
 QY 411 LCEFTDIAKAAGYVTTPIVWSN 434  
 DB 577 LVTFDIAAIKEAGYVVTPIVVTN 600

RESULT 14  
 PTSS-STAXY  
 ID PTSS-STAXY STANDARD; PRT; 480 AA.  
 AC P51184;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE PTS system, sucrose-specific IIBC component (EIIBC-Scr) (Sucrose-  
 DE permease IIBC component) (Phosphotransferase enzyme II, BC component)  
 GN Name=scrA;  
 GN EC 2.7.1.69 (EI-Scr).  
 OS Staphylococcus xylosum.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1288;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 20267 / Isolate C2A;  
 RX MEDLINE=94049686; PubMed=8232209;  
 RA Wagner E., Goetz F., Brueckner R.;  
 RT "Cloning and characterization of the scrA gene encoding the sucrose-  
 RT specific Enzyme II of the phosphotransferase system from  
 RT Staphylococcus xylosum";  
 RL Mol. Gen. Genet. 241:33-41(1993).  
 CC -!- FUNCTION: This is a component of the phosphoenolpyruvate-dependent  
 CC sugar phosphotransferase system (PTS), a major carbohydrate active  
 CC -transport system. The IID domains contain the sugar binding site  
 CC and the transmembrane channel; the IIA domain contains the primary  
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its  
 CC phosphoryl group to the IIB domain which finally transfers it to  
 CC the sugar.  
 CC -!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein  
 CC histidine + sugar phosphate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -!- SIMILARITY: Contains 1 PTS EIIB domain.  
 CC -!- SIMILARITY: Contains 1 PTS EIIC domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X69800; CAA49461.1; -.  
 CC FIR; S39978; S39978.  
 CC InterPro; IPR001996; Ptrans\_EIIB.  
 CC InterPro; IPR003352; Ptrans\_EIIC.  
 CC InterPro; IPR010973; PTS\_II-BC\_sucr.  
 CC Pfam; PF00367; PTS\_EIIB; 1.  
 CC Pfam; PF02378; PTS\_EIIC; 1.  
 CC ProDom; PD001476; PTS\_EIIB; 1.  
 CC TIGRFAMs; TIGR00826; EIIB\_glc; 1.  
 CC TIGRFAMs; TIGR01995; PTS-II-ABC\_beta; 1.  
 CC TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.  
 CC TIGRFAMs; TIGR01992; PTS-IIBC-Tre; 1.  
 CC PROSITE; PS01035; PTS\_EIIB\_CYS; 1.  
 CC Inner membrane; Phosphorylation; Phosphotransferase system;  
 KW Sugar transport; Transferase; Transmembrane.

DR PROSITE; PS00371; PTS\_EI1A.1; 1.  
SQ PROSITE; PS01035; PTS\_EI1B\_CYS; 1.  
SQ SEQUENCE 630 AA; 66978 MW; 6F0218011686ADD5 CRC64;  
Query Match 29.4%; Score 695.5; DB 2; Length 630;  
Best local similarity 32.8%; Pred. No. 1.2e-36;  
Matches 152; Conservative 108; Mismatches 182; Indels 21; Gaps 5;  
QY 2 AMVFPSPVNGYDVAATMAAGEMPWLSFLGLDVAQAGYQGTLPVLVWVSWILATIEKFL 61  
Db 185 ALIYPSIVELHDSALDV-----TFFGIPVLMNYTSTVFPIILLAVFAMSYVEKFCNK 236  
QY 62 RLKGTADFLITPVLTLLTGLTFIALCPAMRWGVDLAHGLQGLYDFGPGVGGLLGLV 121  
Db 237 KIHEAVKFNFTPLILLVIVPVTLIIILGPIGVILNGIASVIOEIFFTSPVLAAGI VAGI 296  
QY 122 YSPVITGLHOSPPIEL-ELFNOGGSFIFATASMANIAQGAACLAFFFLAKSEKLGLA 180  
Db 297 WQVLVIFGIHWGIPIIILNNLSVRGEDIKAVAPAVFSQAGAALGVMLRTKNNKKALA 356  
QY 181 GASGVSAVLGITEPAIFGNLRLWPPFIFIGITAAIGGALIALFNKAVAGAGFLGV 240  
Db 357 GSTSITALEGITEPAVVGVTLPKPKPTMAVISAAGVAIVGHYGSVAVAFAPGLLTIP 416  
QY 241 SI---DAPDMVFLVCVAVTFFIAFGAIAIAGLVLRNGSIDPDATAAPVAGTTKAEA 297  
Db 417 IFYPEDGRGFVAFVIAIISFVLA-----AVLTIVVGFKDPVDEDTLSNMSGSENEVR 471  
QY 298 EAPAEFNSDTIIOAPLTGRAIALSSVSDAMFASGKLGSGVAIVPTKGQLYSPVSGKI 357  
Db 472 BDDKKRPSASEIKSPKLGVEVPLTEVDHVFSSGAMGKGVAVRPKEGELVAPINGTVTS 531  
QY 358 APFGSHAPAVRTKAEDGSNDILMHIGEDTVNLNGTHFNPLKQGDVKAAGELLCEFDID 417  
Db 532 LFEYKHAIGITS---DNGTEIFHVGIDTVQLKGEHFTSFIEQGDVAAAGVLLFEFDE 587  
QY 418 AIKAGYEVTTPIVVSNNYKKTGPVNTYGLGLEAGANLLNVAK 460  
Db 588 RITAAGYDVITPVLITNAKQFSNVQTTDKREVTSDDLIIHVIK 630

Search completed: October 29, 2004, 23:14:21  
Job time : 201 secs

FT DOMAIN 1 ? EI1B.  
FT DOMAIN ? 480 EI1C.  
FT MOD RES 26 26 Phosphocysteine (By similarity).  
FT MOD RES 325 325 Phosphohistidine (By similarity).  
SQ SEQUENCE 480 AA; 51326 MW; AB4E1D9785D84E47 CRC64;  
Query Match 29.78%; Score 701; DB 1; Length 480;  
Best local similarity 51.4%; Pred. No. 4.1e-37;  
Matches 146; Conservative 49; Mismatches 81; Indels 8; Gaps 4;  
QY 1 MAMVFPSPVNGYDVAATMAAG-EMPWMSFLGLDVAQAGYQGTLPVLVWVSWILATIEKFL 59  
Db 194 MILVHPELMSAYDPKALEAGKEIPHNNLFGLEINQVQYQGVLPMLVANVILATIEKGL 253  
QY 60 HKRLKGTADFLITPVLTLLTGLTFIALCPAMRWGVDLAHGLQGLYDFGPGVGGLLGLF 119  
Db 254 RKVIPTVLDNLLTPLLAILSTGFTTFSFVGLPRTGLGWSLSDGLTWLYEFGGATGGILFG 313  
QY 120 LVYSPVITGLHOSFPPIELELF---NQGGSFIFATASMANIAQGAACLAFFFLAK-SE 174  
Db 314 LVYAPIVTGMHHSFIALETQLIADSSSTGSGFIPPTATWSNIAQGAALAAAFIIEKNK 373  
QY 175 KLKGLAGASGVSAVLGITEPAIFGNLRLWPPFIFIGITAAIGGALIALFNKAVAGAA 234  
Db 374 KLKGVSAAGVSALLGITEPAFMFGVNLKLRYPFFIGALVSGIGSGAYIAFFKVKALGTA 433  
QY 235 GFLGVWSIDAPD--WMMLVCAVTVFFTAAGAAIAYGLYLVRN 276  
Db 434 GIPGFISISGQNGWLMHYGIAMIAFVAFVGTVALSYRKKYRN 477  
RESULT 15  
Q9KG19 PRELIMINARY; PRT; 630 AA.  
AC Q9KG19;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PTS system, beta-glucoside-specific enzyme II, ABC component.  
GN Name=BH0296;  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331 (2000).  
DR EMBL; AP001508; BAB04015.1; -.  
DR PIR; H83686; H83686.  
DR HSSP; P20166; 1AX3.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008982; F:protein-N(Pi)-phosphohistidine-sugar phospho. .; IEA.  
DR GO; GO:0005351; F:sugar porter activity; IEA.  
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR011055; Dup hybrid motif.  
DR InterPro; IPR001996; Ptrans\_EI1B.  
DR InterPro; IPR003352; Ptrans\_EI1C.  
DR InterPro; IPR001127; PTS\_EI1A.  
DR InterPro; IPR011297; PTS\_II\_ABC\_beta.  
DR Pfam; PF00358; PTS\_EI1A.1; 1.  
DR Pfam; PF00367; PTS\_EI1B.1.  
DR Pfam; PF02378; PTS\_EI1C; 1.  
DR ProDom; PD001476; Ptrans\_EI1B; 1.  
DR ProDom; PD00243; PTS\_EI1A; 1.  
DR TIGRFAMs; TIGR00830; PTBA; 1.  
DR TIGRFAMs; TIGR01995; PTS-II-ABC-beta; 1.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 23:05:49 ; Search time 40 Seconds  
(without alignments)  
1125.736 Million cell updates/sec

Title: US-09-604-231-2  
Perfect score: 2363  
Sequence: 1 MAWVPSLVNGYDVAAATMAA.....IEAGANLLNVAKEAVPAPP 468  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1058.5	44.8	651	S44257	phosphotransferase
2	953.5	40.4	664	B32243	phosphotransferase
3	759.5	32.1	627	H95951	fusion, PTS system
4	701	29.7	480	S33978	scra protein - Sta
5	695.5	29.4	630	H83686	PTS system, beta-g
6	695	29.4	617	AB1167	phosphotransferase
7	674	28.5	633	AC1436	PTS system, beta-g
8	673.5	28.5	632	C83724	PTS system, beta-g
9	673	28.5	632	S68599	phosphotransferase
10	673	28.5	655	C95220	trehalose PTS syst
11	667	28.2	634	AD1078	PTS system, beta-g
12	664	28.1	627	F95200	PTS system IIRBC c
13	660	27.9	627	B98067	phosphotransferase
14	659	27.9	479	F82432	PTS system, sucros
15	659	27.9	480	D90038	PTS system, sucros
16	658	27.8	705	A95084	phosphotransferase
17	657	27.8	609	I40406	beta-glucoside per
18	646	27.3	479	JQ0781	sucrose uptake pro
19	645	27.3	609	T47097	hypothetical prote
20	640	27.1	631	B42603	beta-glucoside-spe
21	628	26.6	617	AC1421	beta-glucoside-spe
22	612	25.9	628	D97073	PTS system, beta-g
23	588.5	24.9	636	D86807	hypothetical prote
24	580	24.5	618	AC1204	phosphotransferase
25	578.5	24.5	612	A97935	hypothetical prote
26	570.5	24.1	612	B95067	hypothetical prote
27	570	24.1	470	C69725	phosphotransferase
28	569.5	24.1	470	H33926	PTS system, trehal
29	567.5	24.0	625	C25977	phosphotransferase

30	509	21.5	640	2	AB1423	beta-glucoside-spe
31	496	21.0	475	2	C89813	hypothetical prote
32	489	20.7	456	2	S62331	phosphotransferase
33	487	20.6	458	2	H83881	PTS system, sucros
34	473	20.0	455	1	WQEBST	phosphotransferase
35	467	19.8	372	2	I39868	sac operon regulat
36	408	17.3	494	2	AG1231	PTS system trehalo
37	406	17.2	459	2	JU0293	levansucrase synth
38	396.5	16.8	494	2	AF1585	PTS system trehalo
39	390	16.5	460	2	A39938	phosphotransferase
40	367	15.5	473	2	A98281	trehalose specific
41	367	15.5	473	2	C65236	phosphotransferase
42	367	15.5	483	2	AI0449	proteins-Npi-phosph
43	360	15.2	473	2	A86122	PTS system enzyme
44	348.5	14.7	681	2	A89781	hypothetical prote
45	345	14.6	665	2	B96970	PTS enzyme II, ABC

ALIGNMENTS

RESULT 1

S44257  
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose specific - Pediococcus pentosae  
C:Species: Pediococcus pentosaceus  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S44257  
R:Leenhouts, K.K.J.; Bolhuis, A.A.; Kok, J.J.; Venema, G.G.  
submitted to the EMBL Data Library, April 1994  
A:Description: The sucrose and raffinose operons of Pediococcus pentosaceus PPE1.0.  
A:Reference number: S44252  
A:Accession: S44257  
A:Molecule type: DNA  
A:Residues: 1-651 <LEE>  
A:Cross-references: UNIPROT:P43470; EMBL:Z32771; NID:G493728; PIDN:CAA83668.1; PID:G47596  
C:Gene: scra  
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
C:Keywords: phosphotransferase  
F:488-640/Domain: phosphotransferase system glucose-specific enzyme II, factor III homolo

Query Match	44.8%;	Score	1058.5;	DB	2;	Length	651;		
Best Local Similarity	46.7%;	Pred. No.	1.3e-64;						
Matches	221;	Conservative	78;	Mismatches	151;	Indels	23;	Gaps	5;
QY	1	MAWVFP	SLVNGYDVAAATMAAGEMPWWSLFGLDVAAGYQGTVPVLVWVSWILATIEKFLH	60					
Db	194	MIWVFP	SLVNGYDVAAATMAAGRWYVWVFLHVAQAGYQGVLPVLGVAFILATLEKPFH	253					
QY	61	KRLKGT	ADFLITPVLTLTLLTGLTFTIATGPAKRWYGDVLAHGLQGLYDFGPGVGLLFG	120					
Db	254	KHKGAF	DFDFTTPWFAIVITGFLTFTIATGPAKRWYGDVLAHGLQGLYDFGPGVGLLFG	313					
QY	121	VYSPVIT	GLHOSPPRIELELF-----NOGGSFIATASMANIAQGAACLVAFVFLAKSEKL	176					
Db	314	LYSAVIT	GLHQTFPALETQLLANVAKTGGSFIPFVSVANIGQGAATLIAIFATKSQK	373					
QY	177	KGLAGAS	VSAVLGTETPAIFGNLRLRWPFVIGTAAIGGALIALFNKAVLGAAGF	236					
Db	374	KALTSSAG	VSALLGITETPAIFGNLKNLKPFPVFAALASGIASAFTGLFHLVSVAMGPASV	433					
QY	237	LGWVSID	APDMVFLVCAVTFETFAAGAAIAYGLYVLRNCSIDPDATAAPVPAGTTKAE	296					
Db	434	IGFIS	TASKSIPAFWLSAVISFVVAFTPTFTY-----AKRTLGDDRDQVKSAPTSTV---	486					
QY	297	AEAPAE	FSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQGLYSPVSGKIV	356					
Db	487	-----	INVNDEIISAPVTGASBSLKVNDQVFSAEIMGKGAALVPSDDQVVPADGVIT	540					
QY	357	VAFPSGH	AFVTRKAEQGSNNVILMHIGFDTVNLNGHTFNPLKKGQDVEKAGELLCEPDI	416					
Db	541	VTYDSH	AYGIKTTA-----GAELIHLGLDVTNLNGEHTTNVQKGVTHVQDILLGTFTDI	596					

QY 417 DAIRKAGYEVYTPIVVSNKYKTPVNTYGLGETEAGANL--LNVAKKEAVPAT 467  
Db 597 AALKAANYDPTVMLIVTNTANYANVERKLKTNVQAGEQLVALLTAPASSSVAT 649

RESULT 2  
B32243  
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus mutans  
C;Species: Streptococcus mutans  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: B32243  
R;Sato, Y.; Poy, F.; Jacobson, G.R.; Kuramitsu, H.K.  
J. Bacteriol. 171, 263-271, 1999

A;Title: Characterization and sequence analysis of the scrA gene encoding enzyme II (scr)  
A;Reference number: A32243; MUID:89123027; PMID:2536656  
A;Accession: B32243  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-664 <SAT>  
A;Cross-references: UNIPROT:P12655; GB:M22711; NID:g153799; PIDN:AAA26971.1; PID:g153801  
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
C;Keywords: phosphoprotein; phosphotransferase  
F;511-664/Domain: phosphotransferase system glucose-specific enzyme II, factor III homol

Query Match 40.4%; Score 953.5; DB 1; Length 664;  
Best Local Similarity 42.6%; Pred. No. 1.9e-57;  
Matches 205; Conservative 89; Mismatches 156; Indels 31; Gaps 9;

QY 1 MAMVFPPLVNGYDVAATMAAGEMP-----MWSLFGLDVAQAGYQGTVLPVLVWSILA 53  
Db 194 MIMVAPGAANIIGLAANAPISKAATIGATGWNIFGLRHVTQASVYQVPLVAVMILLS 253

QY 54 TIEKFLHKLKGTADFLITPVLTLTLTGFLPIAIGPAMRWGVLAHGLQGLDFGPGV 113  
Db 254 ILERFFHKRLPSAVDFTTPLLSVIITGFLTFIVIGPVNKEVSDWLTNGIIVMLDITGFL 313

QY 114 GGLFLGLVYSPITVITGLHSPPIELEL---FNQG---GSFTFATASMANIAQGAACLA 167  
Db 314 GMVFGALYSPVMTGLHQSPAIETQLISAFQNTGHDGDFIVTASMANVAQGAATPAI 373

QY 168 PFLAKSEKLGLAGASVAVLIGITEPAIFGVNLRWPFPIGIGTAAIGALIALFNK 227  
Db 374 YFLTDXKWKWGLSSSGVSALLGITEPALFGVNLKRYRPFPCALITGSASAAAIGLQV 433

QY 228 AVALGAAGFLGVSDIDPMWMLFVCAVVTFFIAGAAIAVGLYLVRENGSIDPDATAP 287  
Db 434 AVSLGSAGFLGLSLKASSIPVYVVCELISFAIAFAVTVYGG-----KTKAVDVFAAEA 488

QY 288 VPAGTTKABAEAPAFBSNDST-----IIQAPLTGEAIALSSVSDAMFASGLSGVAI 340  
Db 489 VEEATEEVQ-EIPEEASANKAQAQVDEVLAPLAGEAVELTSVNDPVFSSEAMGKGIAI 547

QY 341 VPTKQLVSPVSGKIVAPPSGHAFVARTKABDGNVDILMHIGFTVNLNTHNP LKK 400  
Db 548 KPSGNTVAVPDGTQIAPDTGHAYGI--KSDNGA--EILIHIGIDTVSMEGKGPEQVKQ 603

QY 401 QGDVEKAGELLCERPIDAIKAGYEVYTPIVVSNKYKTPVNTY-GLGRIEAGANLVA 459  
Db 604 ADQIKKGDVLGFTFDSKIAEAGLNTMTFIVTNTADVASVETLASSGTVAVGDSLLEV 663

QY 460 K 460  
Db 664 K 664

RESULT 3  
H96951  
PTS system, beta-glucosides specific IIABC component [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: H96951  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: H96951  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-627 <KUR>  
A;Cross-references: UNIPROT:Q9L8G6; GB:AF001437; PIDN:AAK78403.1; PID:g15023277; GSPDB:G15023277  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC0423  
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match 32.1%; Score 759.5; DB 2; Length 627;  
Best Local Similarity 36.7%; Pred. No. 3e-44;  
Matches 163; Conservative 80; Mismatches 176; Indels 25; Gaps 5;

QY 3 MVFPLVNGYDVAATMAAGEMPWSLFGLDVAQAGYQGTVLPVLVWSWILLATIEKFLH 62  
Db 186 MIHPDLQNAW---TLGEGIKHTINIFGLNIGMVGYQGTLPILISVVMWSYIEKGLR 241

QY 63 LKGTADFLITPVLTLTLTGFLTFIAGPAMRWGVLAHGLQGLYDFGPGVGLLFGIV 122  
Db 242 VPEALDILLTPFLTMITIGTFPAMVVIQGGRFVGDLSLGLQTLVNTTGFSGVLFG 301

QY 123 SPIVITGLHOSFPPIELELFNQG---SFIFATASMANIAQGAACLAFFLAKSEK 179  
Db 302 SLIVITGIHSHFAIEAGLLANPAIHKNFLPLPIWSMANVAQGAALAAVYFTRDKM 361

QY 180 AGASGVAVLIGITEPAIFGVNLRWPFPIGIGTAAIGALIALFNKAVAGAGFLGV 239  
Db 362 AAPASFCLLIGITEPAIFGVNLRWPFPIAGALGAAGGVIVFTKVAMTAVGTGIP 421

QY 240 VSIDAPDMWMLFVCAVVTFFIAGAAIAVGLYLVRENGSIDPDATAPFPAGTTKAE 299  
Db 422 AIVQSGFLNVIAMI-----LAFGAFIIMVILGIKEITEEDLNKETWNGDIKVEE 477

QY 300 PAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGLSGVAIVPTKQLVSPVSGKIV 359  
Db 478 -----VVSFVNGKVLKKNVPDKTFAGLIGDGDIGVDPEDGEVVSPIDGTVV 527

QY 360 PSGHAFVARTKABDGNVDILMHIGFTVNLNTHNP LKKQGDVKAAGELLCERPIDAI 419  
Db 528 ETKHAIAMKSK---NGVEMLIHIGIDTVKMEGNGFKSFINDGEEVKKGDKLIQ 583

QY 420 KAAGYEVYTPIVVSNKYKTPVNT 443  
Db 584 KEKAVSPILVITVNHEDMGFVNS 607

RESULT 4  
S39978  
scrA protein - Staphylococcus xyloso  
C;Species: Staphylococcus xyloso  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S39978  
R;Wagner, E.; Goetz, F.; Brueckner, R.  
Mol. Gen. Genet. 241, 33-41, 1993  
A;Title: Cloning and characterization of the scrA gene encoding the sucrose-specific en  
A;Reference number: S39976; MUID:94049686; PMID:8232209  
A;Accession: S39978  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-480 <WAG>  
A;Cross-references: UNIPROT:P51184; EMBL:X69800; NID:g407905; PIDN:CAA49461.1; PID:g4079  
C;Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

Query Match 29.7%; Score 701; DB 2; Length 480;  
Best Local Similarity 51.4%; Pred. No. 2.2e-40;  
Matches 146; Conservative 49; Mismatches 81; Indels 8; Gaps 4;

QY 1 MAMVFPPLVNGYDVAATMAAG-EMPWSLFGLDVAQAGYQGTVLPVLVWSWILLATIEKFL 59

O.; Jones, L.M.; Karst, U.

PTS system, beta-glucosidases specific enzyme IIABC homolog lin0026 [imported] - *Listeria* f  
C/Species: *Listeria innocua*  
C/Date: 27-Nov-2001 #sequence revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AC1436  
R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Enuian, K.D.; Pshi, H.;  
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001  
A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madieno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A;Title: Comparative genomics of *Listeria* species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AC1436  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-633 <GLA>  
A;Cross-references: UNIPROT:Q92FS7; GB:AL592022; PIDN:CAC95259.1; PID:g16412447; GSPDB:C  
A;Experimental source: strain Clp11262  
C;Genetics:  
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
Query Match 28.5%; Score 674; DB 2; Length 633;  
Best Local Similarity 36.2%; Pred. No. 2.1e-38;  
Matches 163; Conservative 75; Mismatches 179; Indels 30; Gaps 10;  
QY 2 AMVPPSL--VNGYDVAATMAAG---EMPMSLF-GLDVAQAGYQGTIVLPVLVSWILATI 55  
DB 181 ALVYPTLAGISAGDPVITLFTAGTIFESPVHTFGLIPVILMSYASSVPIIATIFGSKV 240  
QY 56 EKFLHKLKGTADFLITPVLLTLLTGTLTIAIGPAMRWVGDVLAHGLQGLYDFGGPYGG 115  
DB 241 EKGFKKIIPDVVKTFFVFCVLLVVPITFLVIGPIATWAGQLLCAGTIWYNLSPVVG 300  
QY 116 LFLGLVYSPVITGLHQSPPIEL-ELFNQGGSFIFATASMANIAQGAACLAIVFLAKSE 174  
DB 301 LILGFMQVFFVFLGLHWGLIPVAINNLTGSDPVLAMFAGSFAQIAGVLAIVFFKTRNK 360  
QY 175 KLKGLAGASGVAVLIGTEPAIFGNVRLRPFFIGITGTAIGGALIALENIKAVAGAA 234  
DB 361 KIKLSIPAFISGIFGVTEPAIYVTLPLKPFINSIAGAVGGIIGFVSAKVIMGL 420  
QY 235 GFLGVVSIDAP-----DMVPLVCAVVTFFTAAGAAIAGLYLVRRNGSIDPDATAFVP 289  
DB 421 GFLGLENFPQAGAGTSAPFWVIVIAVISFI-----LGFILTYVAGFKDP--AEAVV 470  
QY 290 AGTTKAAEAPAEFNSDSTIIQAPITGEAIALSSVDSAMFASGKLGSGVAIVPTGQLVS 349  
DB 471 EETNTEGETLIE---RETIPAPVVGRIIVTLADVKDEAFSSGALGKGVAIITPTVGRVA 526  
QY 350 PVSGKIVVAPFGSHAPAVTKAEDGNSVDILMHIGDVTNLTNGTHFNPLKKGDSVKAGE 409  
DB 527 PAAGTVTITFTPTGHAIGITT--NDGA--EVLHIGMDTVOLEKGTPTAHVKQGDVIEKQ 582  
QY 410 LICEFDIDAIKAAGYVVTPIVVSN 434  
DB 583 LUTEFDIEGKAGYDVITPVPVVTN 607  
RESULT 8  
C83724  
PRS system, beta-glucoside-specific enzyme II, ABC component bglP [imported] - Bacillus  
C;Species: Bacillus halodurans  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C;Accession: C83724  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A;Reference number: AB3650; MUID:20512582; PMID:11058132  
A;Accession: C83724  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-636 <STO>  
A;Cross-references: UNIPROT:Q9KF90; GB:AF001509; GB:BA000004; NID:g10173176; PIDN:BA043  
A;Experimental source: strain C-125  
C;Genetics:  
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
Query Match 28.5%; Score 673.5; DB 2; Length 636;  
A;Gene: bglP

Best Local Similarity 34.0%; Pred. No. 2.2e-38;  
Matches 163; Conservative 78; Mismatches 193; Indels 45; Gaps 9;  
QY 2 AMVPPSLVNGYDVAATMAAGEMPMSLF-----GLDVAQAGYQGTIVLPVLV 47  
DB 183 SLVYPTLV-----VLTEGE-PLYTLFTGTIFESPVHTFGLIPVILMSYATSVIPIIL 234  
QY 48 VSWILATIEKFLHKLKGTADFLITPVLLTLLTGTLTIAIGPAMRWVGDVLAHGLQGLY 107  
DB 235 AAYFASKVEARLAKIIPDVVKTFLLVPFFTLVIVPLTFLVIGPIATWAGQLLGQFTLMVY 294  
QY 108 DFGPVGGLLFLVYSPVITGLHQSPPIEL-ELFNQGGSFIFATASMANIAQGAACLA 166  
DB 295 NLSPITAGAFGLGFWQVFFVFLGHWGLIPIAINNLVVGSDPVLAMFASFAQIGAVAA 354  
QY 167 VFPLAKSEKLGLAGASGVSAVLGITEPAIFGNVRLRPFFIGITGTAIGGALIALFNI 226  
DB 355 VWLKIKQKQKWTLSVPAFISGIFGVTEPAIYGVTLPLKRPFIISCIAAVAGGALIGLFRS 414  
QY 227 KAVALGAAGFLGVVSI---DAPDMVMF-LVCAVVTFFIAGAAIAYGLIYVRRNGSIDP 281  
DB 415 QGYIIGLGLIGFIPSLRHPADGMDAGFWGIVIAVVAFAV-LGFLTITLFLGLKSGNASDEQ 473  
QY 282 DATAAPVAGTTKAAEAPAEFNSDSTIIQAPITGEAIALSSVDSAMFASGKLGSGVAIV 341  
DB 474 TETKAHTSTGTGEKEE-----ISSPFNGSVITLSEIKDEAFSSGALGEGIAIE 521  
QY 342 PTKGQLVSPVSGKIVVAPPSGHAFAVTKAEDGNSVDILMHIGDVTNLTNGTHFNPLKQ 401  
DB 522 PSEKGLSPVSGMWTALYPTTHALGIIT---DRGAELLIHIGLDTVQLDGKFFTAHTIQ 577  
QY 402 GDEVKAGELLCEFDIDAIKAAGYVVTPIVVSNNKKTGPVNTYGLGEITAEAGANLNAVAK 460  
DB 578 GAQVEKGDLLITEFDIKEIKAAAGYAVTTPVIVTNHKYQGLFLTDKQQVNAGRDLLELTR 636  
RESULT 9  
S68599  
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus sobri  
N;Alternate names: sucrose-specific enzyme II  
C;Species: Streptococcus sobrinus  
A;Variety: strain 6715  
C;Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
C;Accession: S68599  
R;Chen, Y.Y.M.; Lee, L.N.; LeBlanc, D.J.  
Infect. Immun. 61, 2602-2610, 1993  
A;Title: Sequence analysis of scrA and scrB from Streptococcus sobrinus 6715.  
A;Reference number: S68598; MUID:93273516; PMID:8500898  
A;Accession: S68599  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-632 <CHE>  
A;Cross-references: EMBL:L06791  
C;Genetics:  
A;Gene: scrA  
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
C;Keywords: phosphotransferase; sugar transport system  
F;480-632/Domain: phosphotransferase system glucose-specific enzyme II, factor III homol  
Query Match 28.5%; Score 673; DB 2; Length 632;  
Best Local Similarity 36.4%; Pred. No. 2.4e-38;  
Matches 168; Conservative 76; Mismatches 197; Indels 20; Gaps 8;  
QY 1 MAMVPPSLVNGYDVAATMAAGEMPMSLFGLDVAQAGYQGTIVLPVLVSWILATIEKFLH 60  
DB 191 LMAVNPALPNAYAV---ASGDAKALTTFFGF-IPVVGYQGTIVLPFAFFVGMIGARLENWLH 245  
QY 61 KRLKGTADFLITPVLLTLLTGTLTIAIGPAMRWVGDVLAHGLQGLYDFGPGVGGLLFGL 120  
DB 246 KRVPEALDLTLTPTLFTFLVMSILGLFAIGVFFHVSFVTVLAATEWILALPFGIAGIILGG 305  
QY 121 VYSPVITGLHQSPPIELFLFNQGGSFIF-ATASMANIAQGAACLAIVFLAKSEKLGKL 179

Db 306 LQQVIVTVGVHHIFNLETQLLAETKANPNPILLSAATAGQGVAVLAIVAKTKSAKLKAL 365  
QY 180 AGAGSVAVIGITEPAIFGVNLRWPFIFIGTAAICGALIALFNKIKAVAGAGFLGV 239  
Db 366 AYPALSAAIGITEPAIFGVNLRWPFIFIGTAAICGALIALFNKIKAVAGAGFLGV 425  
QY 240 VSDADPMVMFLYCAVUTPFIAGCAIAYGLYLVRNGSIDDPDATAAPVAGTTKAEAEA 299  
Db 426 LFLPNSQMPYIVSIITVACAIAP--ALTYY--YFGYADKEEDVSARKPEAPAAAPVAETET 481  
QY 300 PAEFSNDSTIIQAPLTGEALSSVSDAMFASGKLGSGVAIVPTKQGLVSPVSGKIIVAF 359  
Db 482 KSE-----VIASPLDGEAVELSKVNDPFSSEAMGKIIVKPSGNTVYSPVNGTVQIAF 535  
QY 360 PSCHAFVATKAEADGNSVDILMHIGFDVTNLNGTHFNPLKQKQDEVKAGBELLCEFDIDAI 419  
Db 536 ETGHAYGL--KSDNGA--EVLHVGIDTVSMNGTGFQDKVAANQTVKGVGDLVLTFFSAKI 591  
QY 420 KAAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLLINAK 460  
Db 592 ABAGLDDTTWIIITNTADYSEVKPLAAGQLAHAGLLELNK 632  
RESULT 10  
C95220  
trehalose PTS system, IIABC components [imported] - Streptococcus pneumoniae (strain TIC  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: C95220  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Bisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
nson, T.; Hickey, E.K.; Holt, I.B.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; PMID:21357209; PMID:11463916  
A:Accession: C95220  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-655 <KUR>  
A:Cross-references: UNIPROT:Q97NW9; GB:AE005672; PIDN:AAK75956.1; PID:g14973388; GSPDB:C  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP1884  
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
Query Match 28.5%; Score 673; DB 2; Length 655;  
Best Local Similarity 35.3%; Pred. No. 2.5e-38;  
Matches 170; Conservative 87; Mismatches 173; Indels 52; Gaps 16;  
QY 1 MAMVPPSLVNGYDVAAATMAAGEMP--MWSLFLGLDVAQAGYQGTLPVLVWSMILATIEKF 58  
Db 195 ICLVSPQLLNAYAVASTPAADIAANWNNFGYFTVNRIGYQAVIPALLAGLSLYLIF 254  
QY 59 LHKRLKGTADFLITPVLILLTGTFTTFAIGPAMRWGVDVLAHQGLYDFG--GPVG-- 114  
Db 255 WHKHIPVISMIFVFPFLSLIFALILAHVTLGP-----IGWTICQGLSSVVLVAGLTGPKVWL 310  
QY 115 -GLLFLGLVSPVITGLHOSPPPIELF--NOGGSFIFATASMANIAQGAACLAFFLAK 172  
Db 311 FGAIFGALYAPFVITGLHMTNADTQLIADAGGTALWPMIALSNIAQGSVAFFAYFMR 370  
QY 173 -SEKLKLAGASGVSALGITPEAIFGVNLRWPFIFIGTAAICGALIALFNKIKAV 231  
Db 371 HDREAQVSLPATISAYLGVTEPALFGVNVKIYFFVAGMTGSALAGMLSTFVNTAASI 430  
QY 232 GAAGFLGVSDIDAPDMWF-----LVCAVUTPFIAGCAIAYGLYLVRNGSIDDPDATAAP 287  
Db 431 GIGGLPGLSLIQPQMLPFPAGTMLVAVVPMMLTF-----FFRKAGLFTK----- 475  
QY 288 VPAGTTKAEAE-----APAEFNSDSTI-----IQAPLTGEALSSVSDAMFASGKLGSG 337  
Db 476 -TEGDTNLQAEFVQAEAEFVNHPELVTSVELISPTGQVKELSQATDPIFASGVMGQG 534

QY 338 VAIVPTKQGLVSPVSGKIIVVAFPSCHAFVATKAEADGNSVDILMHIGFDVTNLNGTHFNP 397  
Db 535 LVTEPSQGEITSPVNGTIVTLFFTKAIGT--VSDEG--VELLIHGMDTVGLDGKGFES 590  
QY 398 LKQKQGEVKAELLCEFDIDAIKAAGYEVTTPIVWSN--YKKTGPVNTYGLGEIEAGAN 454  
Db 591 LVVQGDHVTVGQQLIRFDMVIVKAGLVTFPVIINQDAYTATIP--GTPT--TIQAGAS 648  
QY 455 LL 456  
Db 649 LM 650  
RESULT 11  
AD1078  
PTS system, beta-glucosides specific enzyme IIABC homolog lmo0027 [imported] - Listeria n  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AD1078  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Psihi, H.;  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitournam, A.; Mat  
ok, C.; Schluter, T.; Simoes, N.; Trierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1078  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-634 <GLA>  
A:Cross-references: UNIPROT:Q8YAT6; GB:NC\_003210; PIDN:CAC98242.1; PID:g16409386; GSPDB:C  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0027  
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
Query Match 28.2%; Score 667; DB 2; Length 634;  
Best Local Similarity 34.6%; Pred. No. 6.2e-38;  
Matches 156; Conservative 81; Mismatches 172; Indels 42; Gaps 10;  
QY 2 AMVFPSSL--VNGYDVAAATMAAG--EMPWMSLF--GLDVAQAGYQGTLPVLVWSMILATI 55  
Db 182 ALVYFTLAGITAGDPLYLFTAGTIFESPIHVTFLGIFVILMSYASSVPIIVATYFGSKV 241  
QY 56 EKFLHKLKGTADFLITPVLILLTGTFTTFAIGPAMRWGVDVLAHQGLYDFGPGVGG 115  
Db 242 EKGKKIIPDVIKTFVPVFFCTLLIVVPITFIVGPIATWAGQLLGGATVWVNLSPFAG 301  
QY 116 LLFGLVSPVITGLHOSPPPIEL--ELFNQGSFIFATASMANIAQGAACLAFFLAKSE 174  
Db 302 LTLGFMQVQVIFGLHMLGVPAVAINNLTVLGHDPILAMTFGASFAQIGAVLAVFFKSRNK 361  
QY 175 KLKLAGASGVSALGITPEAIFGVNLRWPFIFIGTAAICGALIALFNKIKAVAGAA 234  
Db 362 KIKLSIPAFISFGVTEPAIYGVTLPLKPPFNINSCIAAGGIGGGIIGFAGSQTYIMGL 421  
QY 235 GFLGVVSDIDAP-----DMVMFLVCAVUTPFIAGCAIAYGL-----YLVNRNGSIDPDA 283  
Db 422 GIFGLNPFKPGSGISGSEFWVVIIVISFILFTVVGFKDPADVVPVQSNTIVE--- 478  
QY 284 TAAPVAGTTKAEAEAPAEFNSDSTIIQAPLTGEALSSVSDAMFASGKLGSGVAIVPT 343  
Db 479 -----GETLIERET-----IPAPVVGIVTLADVKDEAFSSGALGKGVAIPT 521  
QY 344 KGQLVSPVSGKIIVVAFPSCHAFVATKAEADGNSVDILMHIGFDVTNLNGTHFNPLKQGD 403  
Db 522 VGRVVAAGTVTITFTPTGHAIGITTK--DGA--EVLIIHGMDTVQLEKGFTHFVKQGD 577  
QY 404 EVKAGELLCEFDIDAIKAAGYEVTTPIVWSN 434  
Db 578 VIEKQLLTETFDIEGIIKAAGYDVTTTPVVNT 608

RESULT 12  
F95200  
PTS system IIBC components [imported] - Streptococcus pneumoniae (strain TIGR4)  
C/Species: Streptococcus pneumoniae  
C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C/Accession: F95200  
R/Tetlin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A/Authors: Iofcus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A/Reference number: A95000; MUID:21357209; PMID:11463916  
A/Accession: F95200  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-627 <KUR>  
A/Cross-references: UNIPROT:Q97PB8; GB:AE005672; PIDN:AAK75799.1; PID:gl4973217; GSPDB:C  
A/Experimental source: strain TIGR4  
C/Genetics:  
A/Gene: SPI722  
C/Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
Query Match 28.1%; Score 664; DB 2; Length 627;  
Best Local Similarity 34.3%; Pred. No. 9.8e-38;  
Matches 159; Conservative 92; Mismatches 187; Indels 26; Gaps 10;  
QY 1 MAMVPSLVNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTVPVPLVVSMTLATEIKFLH 60  
DB 182 MMLVSGSLPNAWAVA---QGEVETAMNFFGF-IPVVGLOGSVLPFAFIIGVVGAKFEKAVR 237  
QY 61 KRLKGTADFLITPVLTLTLLTGFTLTFIAIGPAMRWGVDVLAHGLQGLYDFGVPVGLLFLGL 120  
DB 238 KVPDVIDLLVTPFVTLVMSILGLFVIGPVFHVVENVILIATKAILSMPPFLGGLIGG 297  
QY 121 VYSPVITGLHQSPPIELFNQGGSPF-ATASMANIAGAACLAIVFFLAKSEKLGK 179  
DB 238 VHQLIVSGVHHIFNLLEVLAAADHANFPNFIITAAATAQGAATVAVGVTKNPKLKL 357  
QY 180 AGAGSVNAVIGTETPAIGVNLRLRWPFFIGIGTAAIGGALIALFNKAKALGAAGFLG- 238  
DB 358 APPAALSFAFLGITEPAIGVNLRLRWPFFLFLIAGAGGLASILGLAGTNGITIIIPGT 417  
QY 239 VVSDADPMWFLVCAVVTFFIAFGAATAYGLYLVRRNGSIDPDATAAPVPAGTTKARAE 298  
DB 418 MLYVNGQLPQYLLWVAVSFALGTYMFGY-----EDEVDTAAAKRAEVAEKEE 470  
QY 299 -APAEFNSDSTIIQAPLTGEATLSVSDAMFASGLSGVAIVPTKQLVSPVSGKIIV 357  
DB 471 VAPAAALQNETLV--TPVIGDVVALADVNDPVFSSGAMGGIIVKPSQGVVYAPADAESVI 528  
QY 358 APPSGHAFVAVTRKAEDGSNVDILMHIGFTVNLNTHFNPLKKQGVKAGELLCEFFDID 417  
DB 529 APPTGHAFGLKTR----NGAEVLHVGIDITVSMNGDGFTTKAQNKKVAGDVLGTFFDSN 584  
QY 418 AIKAAGYEVTTPIVWSN---YKKTGPVNTYGLGEIETAGANLLNV 458  
DB 585 KIAAAGLDDTTWIVTNTGDYASVAPVAT---GSVAKGDVAVIEV 625  
QY 299 -APAEFNSDSTIIQAPLTGEATLSVSDAMFASGLSGVAIVPTKQLVSPVSGKIIV 357  
DB 471 VAPAAALQNETLV--TPVIGDVVALADVNDPVFSSGAMGGIIVKPSQGVVYAPADAESVI 528  
QY 358 APPSGHAFVAVTRKAEDGSNVDILMHIGFTVNLNTHFNPLKKQGVKAGELLCEFFDID 417  
DB 529 APPTGHAFGLKTR----NGAEVLHVGIDITVSMNGDGFTTKAQNKKVAGDVLGTFFDSN 584  
QY 418 AIKAAGYEVTTPIVWSN---YKKTGPVNTYGLGEIETAGANLLNV 458  
DB 585 KIAAAGLDDTTWIVTNTGDYASVAPVAT---GSVAKGDVAVIEV 625  
RESULT 13  
E98067  
phosphotransferase system enzyme II (EC 2.7.1.69) scra [imported] - Streptococcus pneum  
C/Species: Streptococcus pneumoniae  
C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C/Accession: E98067  
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.O.; Lu, J.; Matsushima, P.; McAhren, S.; M  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A/Reference number: A97872; MUID:21429245; PMID:11544234  
A/Accession: E98067  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-627 <KUR>  
A/Cross-references: UNIPROT:Q8DNN8; GB:AE007317; PIDN:AAL00370.1; PID:gl5459232; GSPDB:G  
C/Genetics:  
A/Gene: scra  
C/Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase  
C/Keywords: phosphotransferase  
Query Match 27.9%; Score 660; DB 2; Length 627;  
Best Local Similarity 34.3%; Pred. No. 1.8e-37;  
Matches 159; Conservative 187; Indels 26; Gaps 10;  
QY 1 MAMVPSLVNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTVPVPLVVSMTLATEIKFLH 60  
DB 182 MMLVSGSLPNAWAVA---QGEVETAMNFFGF-IPVVGLOGSVLPFAFIIGVVGAKFEKAVR 237  
QY 61 KRLKGTADFLITPVLTLTLLTGFTLTFIAIGPAMRWGVDVLAHGLQGLYDFGVPVGLLFLGL 120  
DB 238 KVPDVIDLLVTPFVTLVMSILGLFVIGPVFHVVENVILIATKAILSMPPFLGGLIGG 297  
QY 121 VYSPVITGLHQSPPIELFNQGGSPF-ATASMANIAGAACLAIVFFLAKSEKLGK 179  
DB 298 VHQLIVSGVHHIFNLLEVLAAADHANFPNFIITAAATAQGAATVAVGVTKNPKLKL 357  
QY 180 AGAGSVNAVIGTETPAIGVNLRLRWPFFIGIGTAAIGGALIALFNKAKALGAAGFLG- 238  
DB 358 APPAALSFAFLGITEPAIGVNLRLRWPFFLFLIAGAGGLASILGLAGTNGITIIIPGT 417  
QY 239 VVSDADPMWFLVCAVVTFFIAFGAATAYGLYLVRRNGSIDPDATAAPVPAGTTKARAE 298  
DB 418 MLYVNGQLPQYLLWVAVSFALGTYMFGY-----EDEVDTAAAKRAEVAEKEE 470  
QY 299 -APAEFNSDSTIIQAPLTGEATLSVSDAMFASGLSGVAIVPTKQLVSPVSGKIIV 357  
DB 471 VAPAAALQNETLV--TPVIGDVVALADVNDPVFSSGAMGGIIVKPSQGVVYAPADAESVI 528  
QY 358 APPSGHAFVAVTRKAEDGSNVDILMHIGFTVNLNTHFNPLKKQGVKAGELLCEFFDID 417  
DB 529 APPTGHAFGLKTR----NGAEVLHVGIDITVSMNGDGFTTKAQNKKVAGDVLGTFFDSN 584  
QY 418 AIKAAGYEVTTPIVWSN---YKKTGPVNTYGLGEIETAGANLLNV 458  
DB 585 KIAAAGLDDTTWIVTNTGDYASVAPVAT---GSVSKGDVAVIEV 625  
RESULT 14  
F82432  
PTS system, sucrose-specific IIBC component VCA0653 [imported] - Vibrio cholerae (strain  
C/Species: Vibrio cholerae  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C/Accession: F82432  
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; P  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A/Reference number: A82035; MUID:20406833; PMID:10952301  
A/Accession: F82432  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-479 <HEI>  
A/Cross-references: UNIPROT:Q9KLT8; GB:AE004395; GB:AE003853; NID:99658068; PIDN:AAF9655  
A/Experimental source: serogroup O1; strain N16961; biotype El Tor  
C/Genetics:  
A/Gene: VCA0653  
A/Map position: 2  
C/Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II  
Query Match 27.9%; Score 659; DB 2; Length 479;

Search completed: October 29, 2004, 23:15:05  
Job time : 42 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 23:14:30 ; Search time 130 Seconds  
(without alignments)  
1167.177 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAMVFPVLVNGYDVAAATWAA.....IEAGANLLNVAKEAVPATP 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10G\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10H\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US10I\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US10J\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2363	100.0	468	15	US-10-450-055-10
2	2363	100.0	661	9	US-09-738-626-6404
3	695	29.4	617	15	US-10-282-122A-60875
4	673	28.5	655	16	US-10-474-776-375
5	660	27.9	627	9	US-09-815-242-13467
6	585	24.8	381	16	US-10-474-776-367
7	578.5	24.5	612	15	US-10-282-122A-73843
8	570.5	24.1	612	16	US-10-474-776-253
9	564.5	23.9	620	15	US-10-282-122A-74439
10	492	20.8	379	15	US-10-282-122A-70209
11	468.5	19.8	683	9	US-09-738-626-6961
12	468.5	19.8	683	15	US-10-450-055-26
13	348.5	14.7	681	15	US-10-282-122A-43922

14	345.5	14.6	679	9	US-09-815-242-5658	Sequence 5658, A
15	345.5	14.6	681	9	US-09-815-242-12270	Sequence 12270, A
16	345	14.6	665	15	US-10-282-122A-51493	Sequence 51493, A
17	338.5	14.3	484	9	US-09-815-242-10809	Sequence 10809, A
18	338.5	14.3	484	15	US-10-282-122A-56917	Sequence 56917, A
19	336.5	14.2	474	15	US-10-282-122A-70588	Sequence 70588, A
20	331.5	14.0	453	15	US-10-282-122A-55853	Sequence 55853, A
21	329.5	13.9	526	15	US-10-282-122A-57960	Sequence 57960, A
22	323	13.7	451	15	US-10-282-122A-57628	Sequence 57628, A
23	323	13.7	484	9	US-09-815-242-12272	Sequence 12272, A
24	323	13.7	484	15	US-10-282-122A-44138	Sequence 44138, A
25	322	13.6	650	15	US-10-282-122A-75030	Sequence 75030, A
26	321.5	13.6	453	15	US-10-282-122A-75796	Sequence 75796, A
27	319.5	13.5	429	15	US-10-282-122A-73368	Sequence 73368, A
28	319.5	13.5	650	15	US-10-282-122A-75740	Sequence 75740, A
29	318.5	13.5	651	15	US-10-282-122A-59437	Sequence 59437, A
30	316.5	13.4	466	15	US-10-282-122A-72912	Sequence 72912, A
31	315.5	13.4	677	15	US-10-282-122A-70806	Sequence 70806, A
32	308.5	13.1	657	15	US-10-282-122A-71894	Sequence 71894, A
33	306.5	13.0	455	9	US-09-815-242-13794	Sequence 13794, A
34	305.5	12.9	454	15	US-10-282-122A-45789	Sequence 45789, A
35	304.5	12.9	687	9	US-09-815-242-5822	Sequence 5822, A
36	300	12.7	719	9	US-09-815-242-12942	Sequence 12942, A
37	300	12.7	450	15	US-10-282-122A-71596	Sequence 71596, A
38	299.5	12.7	676	15	US-10-282-122A-52557	Sequence 52557, A
39	299	12.7	726	16	US-10-474-776-280	Sequence 280, App
40	299	12.7	648	9	US-09-741-669-330	Sequence 330, App
41	298.5	12.6	648	9	US-09-815-242-10089	Sequence 10089, A
42	298.5	12.6	648	15	US-10-282-122A-42726	Sequence 42726, A
43	298.5	12.6	665	15	US-10-282-122A-57245	Sequence 57245, A
44	298	12.6	454	15	US-10-282-122A-45636	Sequence 45636, A
45	293.5	12.4				

ALIGNMENTS

RESULT 1  
US-10-450-055-10  
; Sequence 10, Application US/10450055  
; Publication No. US20040043953A1  
; GENERAL INFORMATION:  
; APPLICANT: BASF Aktiengesellschaft  
; TITLE OF INVENTION: NO. US20040043953A1el genes of Corynebacterium  
; FILE REFERENCE: 936 2000  
; CURRENT APPLICATION NUMBER: US/10/450,055  
; CURRENT FILING DATE: 2003-06-10  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Vers. 2.0  
; SEQ ID NO 10  
; LENGTH: 468  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-10-450-055-10

Query Match	100.0%	Score	2363	DB	15	Length	468
Best Local Similarity	100.0%	Pred. No.	1e-195				
Matches	468	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MAMVFPVLVNGYDVAAATWAAAGEMPWLSFLGDLVDAQAGYQGTVLVPLVVSVLATIEKFLH	60				
Db	1	MAMVFPVLVNGYDVAAATWAAAGEMPWLSFLGDLVDAQAGYQGTVLVPLVVSVLATIEKFLH	60				
QY	61	KRLKGTADFLITPVLITLLTGFTFFIAGPMRWVGDVLAHGLQGLYDFGPGVGLLFLGL	120				
Db	61	KRLKGTADFLITPVLITLLTGFTFFIAGPMRWVGDVLAHGLQGLYDFGPGVGLLFLGL	120				
QY	121	VYSPIVITGLHOSPPPELELEFNQGGSFIFATASMANIAQGAACIAVFFIAKSKLKGIA	180				
Db	121	VYSPIVITGLHOSPPPELELEFNQGGSFIFATASMANIAQGAACIAVFFIAKSKLKGIA	180				
QY	181	GASGVSAVLGITEPAITFGVNLRLFWPFFIGTGAAGALITALENIKAVAGAGFLGVV	240				
Db	181	GASGVSAVLGITEPAITFGVNLRLFWPFFIGTGAAGALITALENIKAVAGAGFLGVV	240				

Db	181	GASGSAVLGITEPAIFGVNLRWPFPIGTAIGGALLIALFNKIKAVAGLGAAGFLGVV	240
Qy	241	SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP	300
Db	241	SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP	300
Qy	301	AESFNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIVVAFP	360
Db	301	AESFNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIVVAFP	360
Qy	361	SGHAFVRTKAEQSGNVNDILMHIGFDTVNLNGTHFNFLKQGGDEVKAGELLCEFDIDAIAK	420
Db	361	SGHAFVRTKAEQSGNVNDILMHIGFDTVNLNGTHFNFLKQGGDEVKAGELLCEFDIDAIAK	420
Qy	421	AAGYEVTTPVSNYKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP	468
Db	421	AAGYEVTTPVSNYKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP	468
RESULT 2			
US-09-738-626-6404			
; Sequence 6404, Application US/09738626			
; Publication No. US20020197605A1			
; GENERAL INFORMATION:			
; APPLICANT: NAKAGAWA, SATOSHI			
; APPLICANT: MIZOGUCHI, HIROSHI			
; APPLICANT: ANDO, SEIKO			
; APPLICANT: HAYASHI, MIKIRO			
; APPLICANT: OCHIAI, KEIKO			
; APPLICANT: YOKOI, HARUHIKO			
; APPLICANT: TATEISHI, NAKO			
; APPLICANT: SENO, AKIHIRO			
; APPLICANT: IKEDA, MASATO			
; APPLICANT: OZAKI, AKIO			
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES			
; FILE REFERENCE: 249-125			
; CURRENT APPLICATION NUMBER: US/09/738,626			
; CURRENT FILING DATE: 2000-12-18			
; PRIOR FILING DATE: 1999-12-16			
; PRIOR FILING DATE: 1999-12-16			
; PRIOR FILING DATE: 2000-04-07			
; PRIOR FILING DATE: 2000-08-03			
; NUMBER OF SEQ ID NOS: 7059			
; SOFTWARE: Patent in ver. 3.0			
; SEQ ID NO 6404			
; LENGTH: 661			
; TYPE: PPT			
; ORGANISM: Corynebacterium glutamicum			
US-09-738-626-6404			
Query Match 100.0%; Score 2363; DB 9; Length 661;			
Best Local Similarity 100.0%; Pred. No. 1.7e-195;			
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MANPPSLVNGYDVAAATWAAGEMPWMSLFGLDVAQAGYQGTVPVLVWSWILATIEKFLH	60
Db	194	MANPPSLVNGYDVAAATWAAGEMPWMSLFGLDVAQAGYQGTVPVLVWSWILATIEKFLH	253
Qy	61	KELKGTADFLITPVLTLLTGTLTATGPMRWGDVLAHGLQGLYDFGPGVGLLFG	120
Db	254	KELKGTADFLITPVLTLLTGTLTATGPMRWGDVLAHGLQGLYDFGPGVGLLFG	313
Qy	121	VYSPVITGLHQSFPPIELEFNQGSPTFATASMANIAGAACLVFFLAKSEKLGIA	180
Db	314	VYSPVITGLHQSFPPIELEFNQGSPTFATASMANIAGAACLVFFLAKSEKLGIA	373
Qy	181	GASGSAVLGITEPAIFGVNLRWPFPIGTAIGGALLIALFNKIKAVAGLGAAGFLGVV	240
Db	374	GASGSAVLGITEPAIFGVNLRWPFPIGTAIGGALLIALFNKIKAVAGLGAAGFLGVV	433
Qy	241	SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP	300

RESULT 3

US-10-282-122A-60875  
; Sequence 60875, Application US/10282122A  
; Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 60875  
LENGTH: 617  
TYPE: PPT

ORGANISM: listeria monocytogenes

US-10-282-122A-60875

Query Match 29.4%; Score 695; DB 15; Length 617;

Best Local Similarity 34.8%; Pred. No. 3.2e-51;

Matches 154; Conservative 82; Mismatches 162; Indels 44; Gaps 6;

Qy 2 ANVPSLVNGYDVAAATWAAGEMPWMSLFGLDVAQAGYQGTVPVLVWSWILATIEKFLH 61

Db 185 ALVYPTMNLNFGAHIIFLOIP-----VLMYSFSFPIILAVNLSLIERFLNS 236

```

QY 288 VPAGTTKAEAE---APAEFSDSTI-----IQAPLTGEAIALSSVSDAMFASGKLSSG 337
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
476 -TEGDTNIQAQFVAQBEAEFVNHEBVELTSTVELIISPLTQGVKELSQATDPIFASGVWGQC 534
QY 338 VAIVPTKGQLVSPVSGKIVVAFPPSHAFVRITKAEDGNSVDILMIHIGFTVNLNGTHFNP 397
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
535 LVIEPSQCELTSPVNGTIVTLFPTKHAIGI--VSDEG--VELLIHIGMDTVGLDGGKGFES 590
QY 398 LKKQGDVEYKAGELLCEPDIDAIIKAAGYEVTPPIVSVN---YKKTGPVNTYGLGETEAGAN 454
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
591 LVVQGDHVTVGQQLIRFDMVKAAGLVTETPVIITNQDAYTATIP-GTYPT-TIQAGAS 648
QY 455 IL 456
Db :
649 LM 650

RESULT 5
US-09-815-242-13467
; Sequence 13467, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13467
; LENGTH: 627
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13467

```

```
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13467
; LENGTH: 627
; TYPE: PRN
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13467

Query Match      27.9%; Score 660; DB 9; Length 627;
Best Local Similarity 34.3%; Pred. No. 3.5e-48;
Matches 159; Conservative 92; Mismatches 187; Indels 26; Gaps 10;

QY      1  MAMVFFSLNVGYDVAAATMAAGEMPMSLFGLDVAQAQYQGTVLPVLVVSVILATIEKFLH   60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      182  MLIVSGSLPNAWAVA---QQGEVTANNFFGF-IDPVGLQGSVLIDPAFTIGVGAKFEKAVR   237
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      61  KRUKGTADFLLITPVLILLTGFUTFAIGPAMRWGDVLAHGHLQGLDYDFGCPVGGLLFGL    120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     238  KVPDPVDIDLVTPEFVTLVNMSILGLFVIGDVFHVVENYIILATKAILSMFPGLGSLFGLG    297
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     121  VYSPIVITGLHQSFPPLEELFNQGSFFIF-ATASMANIAOGACLAVFFLAKSEKUKGL    179
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     298  VHQLIVSUGVHHILNFULEVOLLAADHANPFNAITTAAMTQGGAAATVAVGVKTRNPKLKTL   357
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```





Db 462 SQVKSGLSTKQT-LVAPMTGEMFLFSEVDETFSSKLLGEGFALLPSEGEVYAPDFGEVI 520  
QY 357 VAPPGHAFARTKADGSGNDILMHIGFTVNLNTHNPLKQGDVYKAGELLCEFDI 416  
Db 521 TFPPTKHAVALK-----NTRGVEVLHVIGIDTVELKQGFQOLVSVGVVVRGQALLKMDI 576  
QY 417 DAIKAAGYEVTTPIVYSN 434  
Db 577 DFITSXGYSLISEVVVTN 594

RESULT 10  
US-10-282-122A-70209  
; Sequence 70209, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 70209  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-282-122A-70209

Query Match 20.8%; Score 492; DB 15; Length 379;  
Best Local Similarity 37.3%; Pred. No. 5.9e-34;  
Matches 100; Conservative 63; Mismatches 89; Indels 16; Gaps 5;  
QY 1 MAMVFPVSLVGVYDVAAATMAAGEMPMWSLFGLDVAQAGYQGTGTVLPLVSVSWILATIEKFLH 60  
Db 98 LILMHPQLVSQXD-----LAKGNIPWNLFGLBIKQLNQYQGVLPVLIAAYVLAKIEKGLN 153  
QY 61 KRLKGTADFLITPVLTLLITGLTFTIAIGPAMRWGVCDVLAHGLQGLYDFGPGVGLLFL 120  
Db 154 KVVDHSIKRLVGVVPVALLVTGFLAIIIGFVALLIGTGITSGVTTFIQHAGMIGGAIYGL 213

QY 121 VYSPVITGLHQSFPPIELELEFNQ--GGSFIFATASMANIAQGAACLAIVFFLAKSEKL-- 176  
Db 214 LVAPLVIITGLHMFVLAQDQMLGSSLGSTYLPFIVAISNICGSAAFGAWFYVKKRQVVK 273  
QY 177 -KGLAGASVSAVLGTEPAIGVNLRLRWPFIFIGTAAIGGALIALFNKAVALGAAG 235  
Db 274 EREGALTSCISGMLGVTEPAMFGVNLPLKYPFTTAAISTSCVLGAIYGMNVN----LKGVG 329  
QY 236 FLGV--VSIDAPDMVMFLVCAVVTFFI 260  
Db 330 VGVPAFISIQKEFWFVYLIATAIAIV 357

RESULT 11  
US-09-738-626-6961  
; Sequence 6961, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIALI, KEIKO  
; APPLICANT: YOKOL, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 6961  
; LENGTH: 683  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6961

Query Match 19.8%; Score 468.5; DB 9; Length 683;  
Best Local Similarity 29.8%; Pred. No. 1.5e-31;  
Matches 141; Conservative 79; Mismatches 206; Indels 47; Gaps 9;  
QY 27 SLFGLDVAQAGYQGTGTVLPLVSVSWILATIEKFLHKLKGTADFLITPVLTLLTGLTFLFI 86  
Db 218 TVFGLPMVLYNDYSGQVFPPLIAAIGLYVWEKGLKIIPEAVQMVFPVFPFLLMIPATAF 277  
QY 87 AIGPAMRWGVCDVLAHGLQGLYDFGPGVGLFGLVYSIVITGLHQSFPPIELELEFNQ-G 145  
Db 278 LLGPPFGVGVNGISNLEAINNFPFLLSVIPLLYPELVPLGLHWPVNLAINQINILG 337  
QY 146 GSFIFATASMANIAQGAACLAIVFFLAKSEKLKLAGAS--GVSAVL--GITEPAIFGVNL 201  
Db 338 YDFIQGPMGAWNFACFGLVTVGVFLLS IKERNKAMQVSLGGMLAGLLGGISEPSLYGVLL 397  
QY 202 RLRFPPFFIGTAAIGGALIALFNKAVAGAGFLGVVSDADPMVFLVCAVVTFFIA 261  
Db 398 RPKTYFPLLPGLAGGIVMGIFDIKAYAFVFTSLTIPAMD--PWLGYTIGIAVAFVVS 455  
QY 262 FGAAATAGLYLVRRNGSID-----PDAT-AAPV----- 288  
Db 456 MFLVALD---YRSNEERDEAPAKVAADKQAEEDLKAENATPAAPVAAAGAGAGAGA 512  
QY 289 PAGTTKKAERAPAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAVVPTKGQV 348  
Db 513 AGAATAVAAKFLAAGEVWDIVSFLGKAIPLSEVPDPIFAAGKLGPGIAIQPTGNTVV 572

Qy	349	SPVSGKIIVAFPSGHAFVR	TKAEDGSNDVILMHIGFDT	VNLNGHFNPLKKGDEVKAG	408
Db	573	APADATVILVOKSGHALRL	-----DSGVEILVHVGLD	TVOLGGEGFTVHVRRQQVKAG	628
Qy	409	ELLCEFDIDAIKAAGYEV	TTPIVSVNYKKTGPVNTY	GLCEIEAGANLNLNVAKK	461
Db	629	DPLITFDADFIRSKDLPL	ITPVVSNAAKFGIEGIP	ADQANSTTTIVKNGK	681

```

RESULT 12
US-10-450-055-26
; Sequence 26, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Antiangewandtschaft
; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936_2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 26
; LENGTH: 683
; TYPE: PR1
; ORGANISM: Corynebacterium glutamicum
US-10-450-055-26

```

Query Match.	19.8%;	Score 468.5;	DB 15;	Length 683;
Best Local Similarity	29.8%;	Pred. No. 1.5e-31;		
Matches 141;	Conservative 79;	Mismatches 206;	Indels 47;	Gaps 9

  

QY	27	SLRGLDVAQAGYOGTVLPVLVVWSMILATIEKFIHKRLKGTADFLITPVLTLLTGFLFTFI	86
Db	218	TVZGLPMVLNDYSGQVFPPLTAATGLYWEKGLKKIPEAVQMVFPVPPFSLLIMPATAF	277
QY	87	AIGPAMRWGDVLHAHQGLYDFGPGVGGLLFGLVYSIVITGLHQSPFPPIELELEFQO	145
Db	278	LLGPFGLGVGNGISNLEBANNSPFFILSVIPLYFLVPLGLHPWLNATMIQNTLGG	337
QY	146	GSIFIFATASWANTAQGAACLAFLPFAKSEKIKGLAGAS--GVSAVL--GITPEAIFGVNL	201
Db	338	YDFIQGPMGAWNFACFLVTVGVFTLLSIKERNKAMRQVSLGGLAGLLGGTISEPSLYGVLL	397
QY	202	RLRWPFPIGTGTAAGGALLATFNKIVALGAAGFLGVVSDIDPMVMFLVCVVTFPIA	261
Db	398	RFXKTYFRLLPGCLAGGIVMGIFDIKAYAFVFTSLITPAMD--PWLGYTTIGIAVAFVS	455
QY	262	FGAAIAYGLVLYRRNGSID-----PDAT--AAPV-----	288
Db	456	MEVLVLALD---YASNEHRDEARAKVAADKQABEDLKAEANATPAAPVAAAGAGAGAGA	512
QY	289	PAGTTKAEAEAPAEAFNSNDSTIIQAPLTGEAIALSSVSDAMFASGLGSGVAIVPTKGOLV	348
Db	513	AAGAATAVAAKPKLAAGEVVDIVSPLEGKAIPLSEVEDPIFAAGLGGGTAIQFTGNTVV	572
QY	349	SPVSGKIVIVAPPGCHAFVATKTAEDGSDNVILMHIGFTVNLNGHTFNPKKQGDYKAG	408
Db	573	APADATVILVQKSGHVAIRL----DSGVEILVHVGLDVTQLGEGGFTVHVVERQQVKAG	628
QY	409	ELLCEFDIDAIKAAGVEVTTPIVVSNYKKTGPNTYGLGEIAGANLLNVAKK	461
Db	629	DLPLTIDAFIRSKDLPIFPVVVSNAAKGETEGIPADQANSTTIVKNGK	681

RESULT 13  
US-10-282-122A-43922  
; Sequence 43922, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl

QY 309 IIOAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQGVSPVSGKIIVAFPSGHAFVR 368  
Db 534 IIVHAPLTGEVTPLSEVPDQVSEKMMGDGIAIKPSQGEVRAPENGKIQWIFPKHAIGL- 592  
QY 369 TKAEDGSNDIILMHIGFDTVNLNGTHFNPLKKGDEVKAGELICEFDIDAIAKAAGYEVTT 428  
Db 593 ---VSDSGLELLIHIGLDTVKLNGEGFTLHVESGQEVKQGDLLINFDDLYIRNHAKS DIT 649  
QY 429 PIVVS 433  
Db 650 PIVT 654

RESULT 14  
US-09-815-242-5658  
; Sequence 5658, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5658  
; LENGTH: 679  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-5658

Query Match 14.6%; Score 345.5; DB 9; Length 679;  
Best Local Similarity 24.2%; Pred. No. 6.3e-21;  
Matches 132; Conservative 68; Mismatches 192; Indels 153; Gaps 17;

QY 27 SLFLGDVAQAGYGGTFLVPLVSVSWILATIEKFLHKLKGTADFL----ITPVLTLTLTGF 82  
Db 123 SILGIPTLTGTGFGGIIIGALAAW---CVNKFNINLPSYLGFFAGKRFVPIM-MATTSF 178  
QY 83 LTFPAIGAMRVGVDLAHLQ---GLYDFGPGVGGILFGLVSPVITGLHQS F--- 134  
Db 179 I--LAFPMALIW--PTIQTGLNAFTGLDSTNGTAVFLGFKRLLIPFGHLHIFHAPF 234  
QY 135 -----PPIELELFNQ-----GSFIPATASMANIAQGAACLAFFFLA 171  
Db 235 WPEFGSKNAAGEIIGHGQRIEIQIREGAHLTAGKFMQGEFPVMMFGLPAAALAIYHSA 294  
QY 172 KSEKILKAGAGVSA-----VLGITPEALFGVNLRLWFFIGITGAAGGALIAFNK 227  
Db 295 KPNKKVWAGLMSAALTSFLATGITEPLEFSFLFVAPLLFFI---HAYLDGLSFUTLYLL 351

QY 228 AVALG---AAGFLGVVSIDA-PDMVMF-----LVCVVTFPI----- 260  
Db 352 DVHLGYTFSGGFDYVLLGLVLPNKQWLIVPGLVYAVIYVFFRFLIVKLYKTPGRE 411  
QY 261 -----AFGAIAIGLY-----LVR----- 274  
Db 412 DKSOQAVTASATELPYAVLEAMGKANI KHLDACITRLRVEVNDKSKVDVGLKDLGASG 471  
QY 275 -----RNGSIDPDATAAPVPAGTTTKAEAEAPAEFNSNDST 308  
Db 472 VLEVGNNMQAIFGPKSDQIKHEMQIMNGOVVENPTTMEDDKDET VVVAEDKSATSELSH 531  
QY 309 IIOAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQGVSPVSGKIIVAFPSGHAFVR 368  
Db 532 IIVHAPLTGEVTPLSEVPDQVSEKMMGDGIAIKPSQGEVRAPENGKIQWIFPKHAIGL- 590  
QY 369 TKAEDGSNDIILMHIGFDTVNLNGTHFNPLKKGDEVKAGELICEFDIDAIAKAAGYEVTT 428  
Db 591 ---VSDSGLELLIHIGLDTVKLNGEGFTLHVESGQEVKQGDLLINFDDLYIRNHAKS DIT 647  
QY 429 PIVVS 433  
Db 648 PIVT 652

RESULT 15  
US-09-815-242-12270  
; Sequence 12270, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12270  
; LENGTH: 681  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-12270

Query Match 14.6%; Score 345.5; DB 9; Length 681;  
Best Local Similarity 24.2%; Pred. No. 6.4e-21;  
Matches 132; Conservative 68; Mismatches 192; Indels 153; Gaps 17;

QY 27 SLFLGDVAQAGYGGTFLVPLVSVSWILATIEKFLHKLKGTADFL----ITPVLTLTLTGF 82  
Db 125 SILGIPTLTGTGFGGIIIGALAAW---CVNKFNINLPSYLGFFAGKRFVPIM-MATTSF 180

Search completed: October 29, 2004, 23:26:24  
Job time : 133 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 23:06:04 ; Search time 39 Seconds  
(without alignments)  
795.816 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAMVPSLVNGYDVAAATWRA.....IEAGANLNLNVAKEAVPAP 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.\*
- 1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/1aa/PCUTUS-COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832.5	35.2	496	US-09-134-000C-5822	Sequence 5822, Ap
2	669	28.3	627	US-09-583-110-3124	Sequence 3124, Ap
3	662	28.0	628	US-09-107-532A-5288	Sequence 5288, Ap
4	655	27.7	655	US-09-583-110-2958	Sequence 2958, Ap
5	626.5	26.5	656	US-09-489-039A-8212	Sequence 8212, Ap
6	609	25.8	670	US-09-134-000C-4606	Sequence 4606, Ap
7	583.5	24.7	583	US-09-107-532A-6811	Sequence 6811, Ap
8	572.5	24.2	612	US-09-583-110-2911	Sequence 2911, Ap
9	569	24.1	634	US-09-543-681A-7019	Sequence 7019, Ap
10	562.5	23.8	631	US-09-107-532A-3902	Sequence 3902, Ap
11	537	22.7	316	US-09-134-000C-5482	Sequence 5482, Ap
12	520.5	22.0	243	US-09-134-001C-4114	Sequence 4114, Ap
13	501.5	21.2	427	US-09-107-532A-5530	Sequence 5530, Ap
14	497.5	21.1	526	US-09-134-000C-4715	Sequence 4715, Ap
15	489	20.7	483	US-09-489-039A-13018	Sequence 13018, A
16	412	17.4	590	US-09-107-532A-6004	Sequence 6004, Ap
17	363.5	15.4	481	US-09-543-681A-5156	Sequence 5156, Ap
18	357.5	15.1	475	US-09-489-039A-8977	Sequence 8977, Ap
19	336.5	14.2	478	US-09-134-001C-5065	Sequence 5065, Ap
20	302.5	12.8	696	US-09-134-001C-4296	Sequence 4296, Ap
21	295	12.5	726	US-09-583-110-2782	Sequence 2782, Ap
22	294.5	12.5	635	US-09-489-039A-13395	Sequence 13395, A
23	291	12.3	656	US-09-107-532A-5083	Sequence 5083, Ap
24	290	12.3	688	US-09-543-681A-5236	Sequence 5236, Ap
25	284	12.0	482	US-09-107-532A-6644	Sequence 6644, Ap
26	283	12.0	585	US-09-107-532A-4366	Sequence 4366, Ap
27	275	11.6	482	US-09-489-039A-9909	Sequence 9909, Ap

28	269.5	11.4	460	US-09-543-681A-5773	Sequence 5773, Ap
29	267	11.3	170	US-09-134-001C-3134	Sequence 3134, Ap
30	266.5	11.3	470	US-09-107-532A-5798	Sequence 5798, Ap
31	260	11.0	527	US-09-634-238-368	Sequence 368, Ap
32	256	10.8	196	US-09-543-681A-7659	Sequence 7659, Ap
33	248	10.5	315	US-09-134-000C-3657	Sequence 3657, Ap
34	244	10.3	172	US-09-489-039A-9162	Sequence 9162, Ap
35	230	9.7	686	US-09-252-991A-18115	Sequence 18115, A
36	223	9.4	479	US-09-489-039A-10571	Sequence 10571, A
37	181.5	7.7	153	US-09-107-532A-4365	Sequence 4365, Ap
38	159	6.7	355	US-09-107-532A-5592	Sequence 5592, Ap
39	157	6.6	448	US-09-583-110-4040	Sequence 4040, Ap
40	151.5	6.4	658	US-09-492-709A-352	Sequence 352, Ap
41	144	6.1	101	US-09-134-000C-3693	Sequence 3693, Ap
42	141.5	6.0	650	US-09-583-110-5074	Sequence 5074, Ap
43	139.5	5.9	451	US-09-107-532A-5352	Sequence 5352, Ap
44	137	5.8	495	US-09-252-991A-19278	Sequence 19278, A
45	136.5	5.8	661	US-09-134-001C-4303	Sequence 4303, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-134-000C-5822  
; Sequence 5822, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5822  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5822

Query Match 35.2%; Score 832.5; DB 4; Length 496;  
Best Local Similarity 39.3%; Pred. No. 2.4e-75;  
Matches 184; Conservative 84; Mismatches 171; Indels 29; Gaps 7;  
QY 1 MAMVPSLVNGYDVAAATWRAAGEMPMSLFLGDLVAAQAGYQGTVPVLVTVSVILATIEKFLH 60  
Db 47 MINTPDL-----GGATEYWNIFGVHVAQTNYAYQVIVLASVILLSILEKTFH 95  
QY 61 KRLKGTADFLITPVLTLTLTGLTFIAIGPANRWGDVLAHGLQGLYDFGPGVGLLFLGL 120  
Db 96 KKLPSIDFTFTPLLSVLIIGLFTVIGPMLLSNGITDAIVWLINATGFIGMGIFGG 155  
QY 121 VYSPITVGLHQSPPIELFLNQ-----CGSFIFATASMANIAQGAACLAFFLAKSE 174  
Db 156 TYSLIVMTGLHQSPPIELFLNQ-----CGSFIFATASMANIAQGAACLAFFLAKSE 215  
QY 175 KLKGIAGSGVSVLGITETPAIFGVNLRWPPFPGITGTAIGGALIALFNKVALGAA 234  
Db 216 KTKSLASAGLSALGITEPALFGVNLKYPFPFFCALIGSGIAAAITGLLVAVSLGSA 275  
QY 235 GFLGVSIDAPDMVWFLVCAVWTFPIAFGAIAAGLVIVRNGSIDPDATAPVPAGTK 294  
Db 276 GFLGLSINATSIFFLLCELISVFTAYFYG--RTRSSSIFAAIAEQTSDVISE 333  
QY 295 AEAAPAEFNDSTI--IQAPLTGEALISVSDAMFASGKLGSVATVPTKGLQVSPVS 352  
Db 334 INTNQIANADEPTTETIVSVPLAGEITTLGSVNDPVSSESIGKGIKPNGTIYSPVD 393  
QY 353 GKIIVAPSPGHFAVPTKAEAGSNDILMHIGFTVNLNGTHFNPLKKQG--DEVKAGEL 410

GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS: 7310  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5288:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 628 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...628  
SEQUENCE DESCRIPTION: SEQ ID NO: 5288:  
US-09-107-532A-5288  
Query Match 28.0%; Score 662; DB 4; Length 628;  
Best Local Similarity 33.5%; Pred. No. 6.1e-58;  
Matches 160; Conservative 89; Mismatches 172; Indels 56; Gaps 10;  
QY 1 MAMVPPSLVNGYDVAATMAAGEMPWMSLFLGLDV--AQAGYQGTVPVLVWVSVILATIEKF 58  
Db 185 MALVVP-----AITAVAGAGEAISFFGIPVILSPSYTSSVIPILAVWVQSKLEPF 236  
QY 59 LHKRLKGTADFLITPVLTLLTGTFTTIAIGPMRWVGDVLAHGLQGL----YDFGGPVG 114  
Db 237 VKKVIPOQLMILLVPLVWVMPVLTFLALGP----IGTVAGNALGGLFNSIYGFSPIVA 292  
QY 115 GLLFGLVVSPVITGLHQSPFPIE-LELFNQSGSFIFATSMANTIAQGAACLAFFFLAKS 173  
Db 293 GLIMGSLMQVFWFMGHGQFVDFIMELNIEQYGFVLMPLLPAILAOGGAALVALRTKD 352  
QY 174 EKLKGLAGASGVSAVLGITEPAIFGVNLRRLRPPFFIGITAAIGGALIALFNKAVA--- 230  
Db 353 TKLRALGISTVTSLSFGITEPTVYGVTLPLKKPFIAACISGGIGGAILGFGVKAPSSSL 412  
QY 231 ---LGAAGFLGVVSDIDAPDMVPLVCVVTFPIAGAAIAYGLYVRRNGSIDPDATAAP 287  
Db 413 VSLILTPTPTINTVDGVESNVTVAIVATGIAFVLTILGDFDEQTQDNLE----- 465  
QY 288 VPAGITKAAEAPAFBSNDSTIIQAPLTGEATLSSVSDAMFASGLGSGVAIVPTKGQL 347

GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al.  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: PATH00-07A  
CURRENT APPLICATION NUMBER: US/09/583,110  
CURRENT FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 3124  
LENGTH: 627  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-3124  
Query Match 28.3%; Score 669; DB 4; Length 627;  
Best Local Similarity 34.5%; Pred. No. 1.2e-58;  
Matches 160; Conservative 92; Mismatches 186; Indels 26; Gaps 10;  
QY 1 MAMVPPSLVNGYDVAATMAAGEMPWMSLFLGLDVAAQAGYQGTVPVLVWVSVILATIEKFLH 60  
Db 182 MNLVSGSLPNAWAVA---QGEVITAMFFGP-IPVVGLOGSVLPFIPIGVGAKFEKAVR 237  
QY 61 KRLKGTADFLITPVLTLLTGTFTTIAIGPMRWVGDVLAHGLQGLYDFGVPVGLLFLGL 120  
Db 238 KVPVDVILLVTPFTVLLVMSILGLFVIGPVHVVENVILIATKAILSIPFLGGLGFLGG 297  
QY 121 VYSPVITGLHQSPFPIELPLNQSGSPF-ATASMANIAQGAACLAFFFLAKSEKLKGL 179  
Db 298 VHQLIWSGVHIFNLLEVLQALADHANFPNAILITAAMTAQGAATVAVGVTKNPKLKITL 357  
QY 180 AGASGVSAVLGITEPAIFGVNLRRLRPPFFIGITAAIGGALIALFNKAVALGAAGFLG- 238  
Db 358 APPAALSAPLGLITEPAIFGVNLRFRKPFPLSLIAGAGGLASILGLAGTNGNITIIPT 417  
QY 239 VVSDIDAPDMVPLVCVVTFPIAGAAIAYGLYVRRNGSIDPDATAAPVPAGTTKARAE 298  
Db 418 MLYVNGQLPQYLLMVAVSFALGFALTVMFGY-----EDEVDATAAKQAEVAEKEE 470  
QY 299 -APAFBSNDSTIIQAPLTGEATLSSVSDAMFASGLGSGVAIVPTKGOLVSPVSKIVV 357  
Db 471 VAPPAALONETLY--TPVIGDVVALADVNDPVFSSGAMGGIAVKPSQGVVYAPADAESVI 528  
QY 358 APPSGHAFAPVTRKAEDEGSNDVILMHIGFTVNLNGTHFNPLKKQGDENVKAGELLCEFDID 417  
Db 529 APPTGHAFGLKTR----NGAEVLIHVGIDTVSMNGDGFKAQVQGNKVGAGDVLGTFDSN 584  
QY 418 AIIKAGAEYVTPPIVSN---YKKTGPVNTYGLGEIAGANLLNV 458  
Db 585 KIAAAGLDTTMVIVTNTADYASVAPVAT---GSVAKGDAVIEV 625  
RESULT 3  
US-09-107-532A-5288  
; Sequence 5288, Application US/09107532A  
; Patent No. 6583275

Db 466 -----NKHANAGEPITSARHTLKSPLTGKVLSEVPDQVSSGVMGKGIADPEVGEI 519  
QY 348 VSPVSGKIVVAPPSGHAFVTRKAEDGSNDVILMHIGFDTVNLNGTHENPLKKGDEVKA 407  
Db 520 VAPADGEITITFTGHAVGITT--TDGA--EILIHGMDTVELNGNGFEILVKQGLVKA 575  
QY 408 GELLCEFDIDAIAKAGYEVVTPVWSNYKTKGPVNTYGLGEIBAGANLLNVAKEAV 464  
Db 576 GOLLIRFDEIATRAAGSVITPVVITN-----TDAFADILELDQKEII 618

## RESULT 4

US-09-583-110-2958  
; Sequence 2958, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 2958  
; LENGTH: 655  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-2958

Query Match 27.7%; Score 655; DB 4; Length 655;  
Best Local Similarity 34.4%; Pred. No. 3.4e-57;  
Matches 165; Conservative 91; Mismatches 177; Indels 46; Gaps 15;  
QY 1 MAWFPFSLVNGYDVAATMAAGEMP--MWSLFGLDVAQAQYQGTGVLVNVSVILATIEKF 58  
Db 195 ICLVSPQLLNAYAVASTPAADIAANWNVNFGYFTVNRIGYQAVQVPAIAGLSLSYEIF 254  
QY 59 LHKRLKGTADFLITPVLTLTGTFTTAIGPAMRWGVDVLAHGLQGLYDFG--GPVG-- 114  
Db 255 WRKHTEPVISMFVFPFLSLIPALILAHTVLGP-----IGWTICQGLSSVVLAGLTGPVKWL 310  
QY 115 -GLLFGLYVSPVITGLHQSFPPIELF--NOGGSFIFATASMANIAQGAACLAFFLAK 172  
Db 311 FGAIFGALYAPFVITGLHHNTNAIDTQLIADAGGTALPMIALSNIAQGSVAFFYFMHR 370  
QY 173 SEKLKG-LAGASGSVAVLGITEPAIFGVNLRWRPFFIGTAAIGGALIALENIKAVAL 231  
Db 371 HDECEAQVSLPATISAYLGVTPEALFGVNVKVIYFVAGMTGSALAGMLSVTFNTAASI 430  
QY 232 GAAGFLGVVSDAPDMVMF-----LVCAVVTFFIAFGAAIAYGLYLVRNG---SIDPAT 284  
Db 431 GIGLPGILSIQPOVMLPFGATMLVAIVVPLLTP-----PFRKAGLFTKLEGDTN 481  
QY 285 AAPVAGTTKAE--AEAPBSNDSTIIQAPLTGEAIALSSVSDAMFASGLSGVAIVTKGQL 342  
Db 482 LQABFVAQEEAEFVSHEVELT--SVEIISPLTGQVKLSQATDPVFASGVMGQGLVIEP 539  
QY 343 TKGQLVSPVSGKIVVAPPSGHAFVTRKAEDGSNDVILMHIGFDTVNLNGTHENPLKKG 402  
Db 540 SQGLTSPVNGTVTVLFTKKAIGT--VSDEG--VELLIHGMDFVLDGKGFSLVVQG 595  
QY 403 DEVKAGELLCFFDIDAIAKAGYEVVTPVWSNYK-----TGPVNTYGLGEIBAGANLL 456  
Db 596 DHVIVGQQLIRFDMVKAAGLVETPVIIINQDAYTATITGTPT-----TIQAGASLM 650

## RESULT 5

US-09-489-039A-8212  
; Sequence 8212, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8212  
; LENGTH: 656  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8212

Query Match 26.5%; Score 626.5; DB 4; Length 656;  
Best Local Similarity 33.5%; Pred. No. 2.6e-54;  
Matches 158; Conservative 82; Mismatches 183; Indels 49; Gaps 11;  
QY 2 AMVFPFSLVNGYDVAATMAAGEMPMSLFGLDVAQAQYQGTGVLVNVSVILATIEKFLKH 61  
Db 219 ALTHPLMIQAPE--ASQAPG--NAVEHFLGIPVTFINYSVPIIILASVWSCWLERKSA 275  
QY 62 RUKGTADFLITPVLTLTGTFTTAIGPAMRWGVDVLAHGLQGLYDFGPGVGGLLFGLV 121  
Db 276 LLPSSMKNFPSAICLAVVPLTFLVIGPVATWLSHLLANGYQFIYAFAPWLAGAVLGAM 335  
QY 122 YSPVITGLHQSFPPIEL--ELFNQGSFIFATASMANIAQGAACLAFFLAKSEKLGIA 180  
Db 336 WQCVIFGLHGLVPLNMNMTVLGHSMLPILPAVIAQVAGVLGFIATRDARQVLA 395  
QY 181 GASGVSAVLGITEPAIFGVNLRWRPFFIGTAAIGGALIALFNKAVAG-----AA 234  
Db 396 GSAPGAGFGITEPAIYGLTFLRPPFFGCVAGAGGATAFSNSYAYSGFLNIFPPA 455  
QY 235 GFLGVVSDAP-----DMVMFLCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAP 287  
Db 456 QMIPPGGIDASVWGGLIGTGVAFLACVLTFPA-----GL-----PRGSAAP 497  
QY 288 VPAGTTKAEAPABFSNDSTIIQAPLTGEAIALSSVSDAMFASGLSGVAIVTKGQL 347  
Db 498 -----GAVTVAPAS-AND---ILAPMSGVIALEQVDPDSTFASGLLGKGVAIIPAVGQV 547  
QY 348 VSPVSGKIVVAPPSGHAFVTRKAEDGSNDVILMHIGFDTVNLNGTHENPLKKGDEVKA 407  
Db 548 IAPFGEVASLFOYKHAIGL-----QSDSGIELLIHVGDITVKLDGVPFTHAVKSGDRVQA 603  
QY 408 GELLCEFDIDAIAKAGYEVVTPVWSNYKTKGPVNTYGLGEIBAGANLLNVA 459  
Db 604 GDLLIEFDRQAILDAGYDLVTPIIISNDDYREIDTVASSAVEAGQPLLSVS 655

## RESULT 6

US-09-134-000C-4606  
; Sequence 4606, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4606  
; LENGTH: 670

```
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4606

Query Match      25.8%; Score 609; DB 4; Length 670;
Best Local Similarity 30.9%; Pred. No. 1.6e-52;
Matches 146; Conservative 85; Mismatches 200; Indels 42; Gaps 9;

QY 2 AMVFPISLVNGYDVA---TWAGEMPMWSLFGLDV-----AQAGYQGTVLPLVVS 49
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ALCVPTTQGSALQAFETATAGAGAAAPNLFGLPAYNTFMGIPWVGANYTSSVPIIFII 277
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50 WILATIEKFLHKLKGTADFLITLVLTLLTGTFTTAFGPMRWVGDVLAHGLQGLYDF 109
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 AFAAQVQKVKRIPEVVOQFLVDFVLLIAPLPGFLVIGPIVSMILDLISAGTALMSF 337
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 GGPVGLLFLGLVSPVITGHSFPPIELELFNQGSFIFATASM-ANTIAQGAACIAVF 168
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 SPALYGLILGFFWQVLVIFGLHWSVFLAIWQVTEGSSQVLTGSAASFAQTAVILAMF 397
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
169 FLAKSEKILGAGASGVAVLGIETPAIFGNLELRWPFPIGIGTAAIGGALIALENKA 228
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 FKLKDKLKLCPAIIISFGVTEPAIYGITLPKWPFYISMIGGAVGGVLYLMINNTA 457
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
229 VALGAAGFLGVVSI-----DAPDMWMLVCAVVTFFIAFGAAIAYCL-YLVRNGSIDPDA 283
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 YTMGLGIFGVNFINGGDASGMIOQFIA-----IALAAVVGFLITFFFWKDNVREEE 511
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284 TAAPVPAGTTKAEABAPAFESNDSTIIQAPITGBAIALSSVSDAMFASGKLGSVAIVPT 343
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
512 VI--IDKTTIKEN-----ITSPVKGRVLSLKNAEDEPAFANGALGNGVVIPT 557
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
344 KGQLVSPVSGKIVAVPSPGHFAVTRKAEDGNSVDILMHIGFDTVNLGTHFENPLKQGD 403
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
558 EGVVAPDGGIVTLFTPKHALGLIS-----DNGTELLHIGHDVTQLEGEFEAPVKQGD 613
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
404 EVKAGELICEFDIDAIIKAAGYEVVTPPIVVSNNYKKTGPVNTYGLGRIEAGANILL 456
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
614 RVKRGQTLVTDLGKIKAGSFQIPIVVTNTADYLDILEVGSNEVTSDDLL 666

RESULT 7
US-09-107-532A-6811
; Sequence 6811, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
```

```
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6811:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...583
; SEQUENCE DESCRIPTION: SEQ ID NO: 6811:
US-09-107-532A-6811

Query Match      24.7%; Score 583.5; DB 4; Length 583;
Best Local Similarity 29.1%; Pred. No. 4.8e-50;
Matches 139; Conservative 100; Mismatches 194; Indels 45; Gaps 11;

QY 1 MAMVFPISLVNGYDVAATWAGEMPMWSLFGLDVAQAGYQGTVLPLVVSMTIATIEKFLH 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
115 ITLVSPQLNAYSVAST-AAADIPFDGFAQVOMIGTQAVIFAMLAGFWLAVLEIFFR 173
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 KRLAGTADFATITPVLTLTLLTGLTFIAIGP-----AMRWVGDVLAHGLQGLYDFGPGVGL 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
174 KYIQSISMIPLVPLFSLPTVLAHVILSPIGWTVGSMISNIVNTGLTSSIXW---LPSA 230
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 LFGLVYSPVITGHSFPPIELELFNQGSF-IPATASMANIAQGAACIAVFFLAK-SE 174
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
231 VFGFLYAPLVTITGLHMMNAIDMQLIADFGSTNLWPMIALSNIAQGSALVLAIIFLHRCNK 290
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 KLKLAGASGVASVILGITEPAIFGNLELRWPFPIGIGTAAIGGALIALENKAVALGAA 234
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 KERQISIPAMISCVLGVTEPAMFNGLNKYVFPFVAAMVWGSLAGMFANLNGVRANAIVG 350
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
235 GFLGVVSDIDAPDMWMLVCAVVTFFIAFGAAIAYGLYLVRRNG---SIDP--DATAAPVP 289
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
351 GLPGILALQAEWTWPFIIAMIIAVIIPGLATIF-----RRQGIILKIDPAVPENADVQ 405
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
290 AGTTKAEAEAPAFESNDSTIIQAPITGBAIA-----LSSVSDAMFASGKL 334
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
406 LQTANGATATPQSFEPVSA-----TGTAVATKETFVAAGTIKEITEVNDPVSQRKM 459
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
335 GSGVAIVTKQLVSPVSGKIIVAPPSGHAFVTRKAEDGNSVDILMHIGFDTVNLGTH 394
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
460 GDGYAVEPSNGKVYAPVNGKVTSVFETKHAIGILS-----NEGLEVIVHWGLDTVELKGYP 515
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
395 FNPLKKQGDVKAGELLCFEDIDAIIKAAGYEVVTPPIVVSNNYKKTGPVNTYGLGRIEAG 452
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
516 FNVFVKEGYLVTPELIIAEMDLPEIQAGKKTDIIIVALTINNEKVAGLSLDQSGLVVRP 573

RESULT 8
US-09-583-110-2911
; Sequence 2911, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
```

```
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2911
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2911

Query Match      24.3%; Score 572.5; DB 4; Length 612;
Best Local Similarity 33.3%; Pred. No. 6.7e-49;
Matches 143; Conservative 81; Mismatches 145; Indels 61; Gaps 12;

QY 27 SLFGLDVAQAGYOGTVLPVLVSVWILATIEKFLHKLKGTADFLIPLVTLTLTGLTFI 86
Db 196 SLFGAPVTASYGSSVVPILIMVLMKYTEKMTAKLTPAVTKSFLOFTLVLLVSSCIALV 255
QY 87 AIGPAMRWGDVLAHGLQGLYDFGGPVGGLFGLVTSPIVITGLHOSFPPIELFLNQGG 146
Db 256 VVGPIGVIVGEGSLNVGMVGAGWLTALIGAIMPFIVMTGNHAFAP-----305
QY 147 SFIFATASMA-----NTAOGAACLAFFFLAKSEKLGAGAGSVSAVL-GIT 192
Db 306 --IFLAASATPDVLILPAMLGSLNLAQGAASMAVALKSKNNNTKQIAFAAGFSALLAGIT 363
QY 193 EPAIFGNLRLRWPFFIGTGAIGGALIALF-----NIKAVALGAAGFLGV---VSIDAP 245
Db 364 EPALGVTLKTKPLF-----AAMIGGGLAGLTSVKAYLFAVPSLIALPQFIYSVDP 419
QY 246 -DMVFLVCAVTVFFIAFGAATAYGLYVRRNGSIDPDATAAPVPAGTTKAEAAEAPFS 304
Db 420 SNIVNALIVAVISVITFVLAIRG-----IDEESS-----SNLEKAGVS 461
QY 305 NDSTIIQPLTGEALSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIIVAPPSSGHA 364
Db 462 NKKMIF-SPISGEIIPLSVDQKTFSDKLGIDGVALIPSEGVYAPFDGKITNIPTKHA 520
QY 365 FAVRTKAEKGSNDVIMHIGFTVNLNGTHFNPLKQGEVAGELLCEFFDIDAKAAGY 424
Db 521 IGL--KSEGG--VELLIHGLTVLKGQGFISHVEEGDVFKNQLIFEMDLNLKTKGY 576
QY 425 EYVTTPIVNSN 434
Db 577 ENTVFVITVN 586

RESULT 9
US-09-543-681A-7019
; Sequence 7019, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7019
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7019

Query Match      24.1%; Score 569; DB 4; Length 634;
Best Local Similarity 32.8%; Pred. No. 1.6e-48;
Matches 150; Conservative 90; Mismatches 161; Indels 56; Gaps 16;

QY 1 NAMVPPSLVNGYDVAATMAAGEMPWSLFGDLVQAQYQGTGTVLPVLVSVWILATIEKFLH 60
Db 186 LAMVSP-----EWTSLVKAQE-PV-EPMHPVLVALVKYSSQLIIPALITVNMWSTIERFIV 237
QY 61 KRLKGTADFLITPVLTLLLTGTLFTIATGPAMRWGDVLAHGLQGLYDFGGFVGGLIFGL 120
```

```
Db 238 RIVPENVKVMPVLLVILYSTPIALIAVGPTVSFAQLIADGVLLIQHTGTFIAIPLVA 297
QY 121 VYSPVITGTHOSFPPIELFLNQGG--SFIFATASMANIAQGAACLAFFFLAKSEKLG 178
Db 298 IYPLVLSIGNHKALSPVSIMLVEQKGFDPPIIRVMALCSNWSQAASLAVSVRTKNKTLKQ 357
QY 179 LAGASGVSAVL-GITEPAIFGNLRLRWPFFIGTGAIGGALIALF-----NIKAVALGA 233
Db 358 LAFSASITAFFGITTEPAMYGVNLKUKPMY-----ACMIGGAIAAGLFAIVKLFKAFYVVT 413
QY 234 AGEFLGV---VSIDAPMMVFLVCAVTVFFIAFGAATAYGLYVRRNGSIDPDATAAPV-- 288
Db 414 PGLSLPMMIS-DTDNQV---VNAITLILASVATFIATLII-----GFDDP--TDDPIRD 463
QY 289 -----PAGTTK-----AEAAEAPAEFSNDSTIIQAPLTGEALSVSDAMFASGKLGSG 337
Db 464 EBENNKQAASNTKKQIANSKLPVG-----LISPLQKTVALSVDNETFASGIMPG 516
QY 338 VAIVPTKGOLVSPVSGKIIVAPPSSGHAFAPVRYKAEKGSNDVIMHIGFTVNLNGTHFN 397
Db 517 MALIPTTGKVIAPADGVVDITFSSGHAIGTLV---NNIEMLIHVIGIDTVNLAGQHFTC 572
QY 398 LKKQGEDEVKAGELLCEFFDIDAKAAGYEVTTPIVNSN 434
Db 573 CVVKGQKVTGDTLAEFFDLDAIIAAGYDPTMTIITN 609

RESULT 10
US-09-107-532A-3902
; Sequence 3902, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3902:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
```





```
; SEQ ID NO 13018
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13018

Query Match      20.7%; Score 489; DB 4; Length 483;
Best Local Similarity 38.5%; Pred.No. 1.3e-40;
Matches 104; Conservative 51; Mismatches 101; Indels 14; Gaps 3;

QY      3 MVPEPSLVNGYDVAA---TMAAGEMPWWSLFLGDLVAQAQYQGTVPVLPVLVWSWILATIEKFL 59
Db      213 LTHPALTNANGVAAGFHTM-----NFFGIEVAMIGYQGTVPFPVLLAVWFMSWEKRL 264

QY      60 HKRLKGTADFLITPVLTLTLTGFTLFTFAIGPAMRWGDVLAHGLQGLYDFGPGVGGLLFG 119
Db      265 RRVIPDALDLITLPFLTVIISGFTIALLLIGPAGRALGDGIFSILSTLIISHAGWLAGLLFG 324

QY      120 LVYSPVITGLHQSPFPPIELELFNQ--GGSPFIATASMANIAQGAACLAVFFFLAKSEKL 176
Db      325 GLYSVIVITGLHHSFHAIEAGLLGNPSIGVNFILPIWAMNVAQGGACFAVWFKTKDAKI 384

QY      177 KGLAGAGSVSAVLGITPFAIFGVNLRLRWPPFICIGTAAIGCALIALFNKIKAVALGAAGF 236
Db      385 KAITLPSAFSAMLGITTEAAIFGINLRFVKPFIALLVGGAGGAWVSMHVYNTAVGLTAI 444

QY      237 LGVVSIDAPDMWMLVCAVVTFFIAFGNAI 266
Db      445 PGMAIVQASSLLNYIIGNAIAFAVAFALSL 474
```

Search completed: October 29, 2004, 23:15:53  
Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 22:57:48 ; Search time 156 Seconds  
(without alignments)  
1076.188 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAWVPSLVNGYDVAATMAA.....IEAGANLLNVAKEAVPATP 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2363	100.0	468	4	AAB66707
2	2363	100.0	468	5	ABG80325
3	2363	100.0	468	5	AAG92850
4	2332	98.7	661	4	AAB69080
5	1821	77.1	362	4	AAB66708
6	832.5	35.2	496	7	ADH87937
7	695	29.4	617	5	ABH47495
8	695	29.4	617	6	ABU32951
9	673	28.5	655	6	ABU02378
10	673	28.5	655	6	ABP81458
11	669	28.3	627	8	ADK46609
12	667	28.2	634	5	ABH49833
13	663.5	28.1	676	5	ABP25654
14	662	28.0	628	7	ADU95661
15	660	27.9	627	4	AAU37874
16	659	27.9	627	6	ABM72990
17	655	27.7	655	8	ADK46443
18	647.5	27.4	674	5	ABP25655
19	632	26.7	620	5	ABP27216
20	628	26.6	617	5	ABH49923
21	626.5	26.5	656	7	ABO61695
22	609	25.8	670	7	ADH86721
23	591.5	25.0	639	5	ABP27215
24	588.5	24.9	636	5	ABH54803
25	585	24.8	381	6	ABP81450

ALIGNMENTS

RESULT 1

AAB66707	583.5	24.7	583	7	ADC97184	E. faeciu
ID	AAB66707	standard; protein; 468 AA.	618	5	ABH47649	Listeria
XX	AAB66707;		622	5	ABP28755	Streptoco
XX			612	6	ABU45919	Protein e
DT	09-APR-2001	(first entry)	612	8	ADK46396	Streptoco
DE	C.Glutamicum phosphoenolpyruvate protein #1.		612	6	ABU00950	S. pneumo
XX	Phosphoenolpyruvate; sugar phosphotransferase system; PTS.		612	6	ABP81336	Streptoco
XX	Corynebacterium glutamicum.		612	7	ADP06734	Bacterial
XX			620	5	ABP26854	Streptoco
PN	WO200102583-A2.		620	6	ABU46515	Protein e
XX	11-JAN-2001.		631	7	ADC94275	E. faeciu
XX			618	8	ADN61668	Corynebac
XX			316	7	ADH87597	Enterococ
XX			334	6	ABU02158	S. pneumo
XX			243	5	ABP39289	Staphyloc
XX			640	5	ABH49939	Listeria
XX			427	7	ADC95903	E. faeciu
XX			526	7	ADH86830	Enterococ
XX			379	6	ABU42285	Protein e
XX			478	6	ABM71603	Staphyloc

Claim 4; Page 101-102; 144pp; English.

The present invention relates to Corynebacterium glutamicum phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The PTS nucleic acids and proteins are useful in the identification of microorganisms which can be used to produce fine chemicals, for modulating fine chemical production in C. glutamicum or related bacteria, the typing or identification of C. glutamicum or related bacteria, as reference points for mapping C. glutamicum genome, and as markers for transformation

Sequence 468 AA;

PT	fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins, cofactors or enzymes used in food, feed, cosmetics or pharmaceutical industries.
PT	
PT	
XX	Claim 18; Page 111-113; 176pp; English.
XX	
XX	The present invention relates to the isolation of Corynebacterium glutamicum metabolic pathway (MP) proteins, and the polynucleotide sequences encoding them. The MP proteins are enzymes involved in the metabolism of molecules important for the normal functioning of cells (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or trehalose). The polynucleotide sequences encoding the MP proteins are useful for producing fine chemicals, particularly organic acids, non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, (un)saturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides and enzymes. The fine chemicals are useful in the food, animal feed, cosmetic or pharmaceutical industries. ABG80321-ABG80343 represent the C. glutamicum MP proteins of the invention
XX	Sequence 468 AA;
SQ	
Query Match	100.0%; Score 2363; DB 5; Length 468;
Best Local Similarity	100.0%; Pred. NO. 1.2e-226;
Matches	468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MAMVFPSSLVNGYDVAATMAAGEMPWMSLFGLDVAQAGYQGTGTLVPLVVSILATIEKFLH 60
DB	1 MAMVFPSSLVNGYDVAATMAAGEMPWMSLFGLDVAQAGYQGTGTLVPLVVSILATIEKFLH 60
QY	61 KRLKGTADFLITPVLITLLTGFTLFIAGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG 120
DB	61 KRLKGTADFLITPVLITLLTGFTLFIAGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG 120
QY	121 VYSPIVITGLHQSPPPIELEFNQGSFIATASMANIAQGAACLAFFLAKSEKLG 180
DB	121 VYSPIVITGLHQSPPPIELEFNQGSFIATASMANIAQGAACLAFFLAKSEKLG 180
QY	181 GASGVSALVIGITEPAIFGVNLRRLRWPFFIGTGTAAIGGALIALFNKAVALGAAGFLGV 240
DB	181 GASGVSALVIGITEPAIFGVNLRRLRWPFFIGTGTAAIGGALIALFNKAVALGAAGFLGV 240
QY	241 SIDAPDMWMLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEAP 300
DB	241 SIDAPDMWMLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEAP 300
QY	301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI 360
DB	301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI 360
QY	361 SGHFAVTKAEDGSNDVILMHIGFTVNLNGTHFNPLKQGDVVKAGELLCEFDIDA 420
DB	361 SGHFAVTKAEDGSNDVILMHIGFTVNLNGTHFNPLKQGDVVKAGELLCEFDIDA 420
QY	421 AAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLINVAKEAVPATP 468
DB	421 AAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLINVAKEAVPATP 468
RESULT 3	
AAG92650	
ID	AAG92650 standard; protein; 661 AA.
XX	
AC	AAG92650;
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum protein fragment SEQ ID NO: 6404.
XX	
KW	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis.
XX	
OS	Corynebacterium glutamicum.
PT	
Query Match	100.0%; Score 2363; DB 4; Length 468;
Best Local Similarity	100.0%; Pred. NO. 1.2e-226;
Matches	468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MAMVFPSSLVNGYDVAATMAAGEMPWMSLFGLDVAQAGYQGTGTLVPLVVSILATIEKFLH 60
DB	1 MAMVFPSSLVNGYDVAATMAAGEMPWMSLFGLDVAQAGYQGTGTLVPLVVSILATIEKFLH 60
QY	61 KRLKGTADFLITPVLITLLTGFTLFIAGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG 120
DB	61 KRLKGTADFLITPVLITLLTGFTLFIAGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG 120
QY	121 VYSPIVITGLHQSPPPIELEFNQGSFIATASMANIAQGAACLAFFLAKSEKLG 180
DB	121 VYSPIVITGLHQSPPPIELEFNQGSFIATASMANIAQGAACLAFFLAKSEKLG 180
QY	181 GASGVSALVIGITEPAIFGVNLRRLRWPFFIGTGTAAIGGALIALFNKAVALGAAGFLGV 240
DB	181 GASGVSALVIGITEPAIFGVNLRRLRWPFFIGTGTAAIGGALIALFNKAVALGAAGFLGV 240
QY	241 SIDAPDMWMLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEAP 300
DB	241 SIDAPDMWMLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPAGTTKAEAP 300
QY	301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI 360
DB	301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI 360
QY	361 SGHFAVTKAEDGSNDVILMHIGFTVNLNGTHFNPLKQGDVVKAGELLCEFDIDA 420
DB	361 SGHFAVTKAEDGSNDVILMHIGFTVNLNGTHFNPLKQGDVVKAGELLCEFDIDA 420
QY	421 AAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLINVAKEAVPATP 468
DB	421 AAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLINVAKEAVPATP 468
RESULT 2	
ABG80325	
ID	ABG80325 standard; protein; 468 AA.
XX	
AC	ABG80325;
XX	
DT	15-NOV-2002 (first entry)
XX	
DE	C. glutamicum metabolic pathway (MP) protein #5.
XX	
KW	Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
KW	cofactor; nucleotide; nucleoside; trehalose; fine chemical production;
KW	organic acid; non-proteinogenic amino acid; purine base; carbohydrate;
KW	pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;
KW	aromatic compound; food industry; animal feed; cosmetic industry;
KW	pharmaceutical industry; enzyme.
XX	
OS	Corynebacterium glutamicum ATCC 13032.
XX	
PN	WO200251231-A1.
XX	
PD	04-JUL-2002.
XX	
PF	22-DEC-2000; 2000WO-EP013143.
XX	
PR	22-DEC-2000; 2000WO-EP013143.
XX	
PA	(BADI ) BASF AG.
XX	
PI	Pompejus M, Kroeger B, Zelder O, Schroeder H;
XX	
DR	WPI; 2002-643289/69.
DR	N-PSDB; ABS65346.
XX	
PT	New metabolic pathway genes of Corynebacterium glutamicum for producing

XX PN EP1108790-A2.  
 XX PD 20-JUN-2001.  
 XX PF 18-DEC-2000; 2000EP-00127688.  
 XX PR 16-DEC-1999; 99JP-00377484.  
 XX PR 07-APR-2000; 2000JP-00159162.  
 XX PR 03-AUG-2000; 2000JP-00280988.  
 XX PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX WPI; 2001-376931/40.  
 XX N-PSDB; AAH67869.  
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
 mutation point of a gene, measuring expression of a gene, analyzing  
 expression profile or pattern of a gene and identifying homologous gene.  
 XX Claim 17; SEQ ID NO 6404; 246pp + Sequence Listing; English.  
 XX The present invention provides a number of nucleotide and protein  
 sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 are useful for identifying the mutation point of a gene derived from a  
 mutant of coryneform bacterium, measuring expression amount and analysing  
 the expression profile or expression pattern of a gene derived from  
 Coryneform bacterium, and identifying a homologue of a gene derived from  
 coryneform bacterium. Coryneform bacteria are useful for producing amino  
 acids, nucleic acids, vitamins, saccharides and organic acids,  
 particularly L-lysine. The present sequence is a protein described in the  
 exemplification of the invention. Note: The sequence data for this patent  
 did not form part of the printed specification, but was obtained in  
 electronic format directly from the European Patent Office  
 Sequence 661 AA;  
 Query Match 100.0%; Score 2363; DB 4; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-226;  
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKPLH 60  
 DB 194 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKPLH 253  
 QY 61 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFL 120  
 DB 254 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFL 313  
 QY 121 VYSPVITVGLHQSPPIELFENQGGSFIFATASMANIAQGAACIAVFFLAKSEKLGKLA 180  
 DB 314 VYSPVITVGLHQSPPIELFENQGGSFIFATASMANIAQGAACIAVFFLAKSEKLGKLA 373  
 QY 181 GASGVSALVIGTEPAIFGNVLRNPPFPGIGTAAIGGALLAFNKAVALCAAGFLGV 240  
 DB 374 GASGVSALVIGTEPAIFGNVLRNPPFPGIGTAAIGGALLAFNKAVALCAAGFLGV 433  
 QY 241 SIDAPDMVFLVCAVTFPIAAGAAIAYGLVLRNNGSIDPDAAPVPAGTTKAEAEAP 300  
 DB 434 SIDAPDMVFLVCAVTFPIAAGAAIAYGLVLRNNGSIDPDAAPVPAGTTKAEAEAP 493  
 QY 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIIVAPP 360  
 DB 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIIVAPP 553  
 QY 361 SGHAFVTRKAEAGSNVDILMHI GFDVTNLTNGTHFNPLKKGDEKAGELLCEFDIDAIAK 420  
 DB 554 SGHAFVTRKAEAGSNVDILMHI GFDVTNLTNGTHFNPLKKGDEKAGELLCEFDIDAIAK 613  
 QY 421 AAGVEVTPPIVVSNNKKTGPVNTYGLGEIAGANLLNVAKEAVPATP 468

DB 614 AAGVEVTPPIVVSNNKKTGPVNTYGLGEIAGANLLNVAKEAVPATP 661

# RESULT 4

AAB69080  
 ID AAB69080 standard; protein; 661 AA.

AC AAB69080;

XX 11-SEP-2003 (revised)  
 DT 20-APR-2001 (first entry)

XX Brevibacterium lactofermentum sucrose PTS enzyme II protein SEQ ID NO:2.

XX Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;  
 KW phosphoenolpyruvate:carboxylate phosphotransferase system; glucose;  
 KW coryneform bacterium; phosphoenolpyruvate-sugar transport system.

XX Corynebacterium glutamicum.

OS Corynebacterium glutamicum.

PN WO200102584-A1.

XX 11-JAN-2001.

XX 30-JUN-2000; 2000WO-JP004348.

PR 02-JUL-1999; 99JP-00189512.

XX (AJIN ) AJINOMOTO CO-INC.

XX Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;

XX WPI; 2001-138150/14.

DR N-PSDB; AAF32543.

XX Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme  
 II obtained by cassette ligation-mediated amplification of downstream  
 domain of coryneform bacterium sucrose gene, with sucrose-binding  
 activity.

XX Claim 1; Page 29-32; 45pp; Japanese.

CC The present sequence represents the Brevibacterium lactofermentum sucrose  
 PTS (phosphoenolpyruvate:carboxylate phosphotransferase system or  
 phosphoenolpyruvate-sugar transport system) enzyme II, which has sucrose-  
 binding activity. A coryneform bacteria produced with the sucrose PTS  
 enzyme II gene can have more efficient sugar uptake, and improved amino-  
 acid and nucleic acid productivity. The sucrose PTS gene and its  
 disrupted gene, such as one without the sucrose PTS function, can be used  
 to produce new breeds of coryneform bacterial strains to uptake sugar  
 more efficiently e.g. glucose only or and sucrose, and can have improved  
 amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to  
 standardise OS field)

XX Sequence 661 AA;

Query Match 98.7%; Score 2332; DB 4; Length 661;

Best Local Similarity 98.5%; Pred. No. 2.6e-223;

Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKPLH 60

DB 194 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKPLH 253

QY 61 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFL 120

DB 254 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFL 313

QY 121 VYSPVITVGLHQSPPIELFENQGGSFIFATASMANIAQGAACIAVFFLAKSEKLGKLA 180

DB 314 VYSPVITVGLHQSPPIELFENQGGSFIFATASMANIAQGAACIAVFFLAKSEKLGKLA 373

QY	181	GASGSAVLGITTEPAIFGVNLRWRPFPIGTAAGGALIALFNLIKAVAGAGFLGVV	240
Db	374	GASGSAVLGITTEPAIFGVNLRWRPFPIGTAAGGALIALFDIKAVAGAGFLGVV	433
QY	241	SIDAPDMVFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP	300
Db	434	SIDAPDMVFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP	493
QY	301	AEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQIVSPVSGKIIVVAPP	360
Db	494	AEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQIVSPVSGKIIVVAPP	553
QY	361	SGHAFVRTKAEDGNSVDILMHIGFDTVNLNGTHFNPLKKGDRVKAGELLCEFDIDAIAK	420
Db	554	SGHAFVRTKAEDGNSVDILMHIGFDTVNLNGTHFNPLKKGDRVKAGELLCEFDIDAIAK	613
QY	421	AAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP	468
Db	614	AAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP	661
RESULT 5			
ID	AAB66708 standard; protein; 362 AA.		
XX	AAB66708;		
AC	AAB66708;		
XX	09-APR-2001 (first entry)		
DT	C.glutamicum phosphoenolpyruvate protein #2.		
DE	Phosphoenolpyruvate; sugar phosphotransferase system; PTS.		
XX	Corynebacterium glutamicum.		
KW	WO200102583-A2.		
OS	11-JAN-2001.		
XX	27-JUN-2000; 2000WO-IB000973.		
PR	01-JUL-1999; 99US-0142691P.		
PR	23-AUG-1999; 99US-0150310P.		
PR	03-SEP-1999; 99DE-01042095.		
PR	03-SEP-1999; 99DE-01042097.		
XX	(BADI ) BASF AG.		
PA	Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;		
PI	WPI; 2001-080989/09.		
DR	Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:		
XX	sugar phosphotransferase system proteins or their portions, useful for		
PT	typing or identifying C. glutamicum or related bacteria, and as markers		
PT	for transformation.		
PS	Claim 6; Page 104-106; 144pp; English.		
XX	The present invention relates to Corynebacterium glutamicum		
CC	phosphoenolpyruvate; sugar phosphotransferase system (PTS) proteins. The		
CC	PTS nucleic acids and proteins are useful in the identification of		
CC	microorganisms which can be used to produce fine chemicals, for		
CC	modulating fine chemical production in C. glutamicum or related bacteria,		
CC	the typing or identification of C. glutamicum or related bacteria, as		
CC	reference points for mapping C. glutamicum genome, and as markers for		
CC	transformation		
XX	Sequence 362 AA;		
QS	Query Match 77.1%; Score 1821; DB 4; Length 362;		
XX	Best Local Similarity 100.0%; Pred. No. 1.1e-172;		
XX	Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	107	YDFGGPVGGLLFGVYSPVITGLHQSFPPIELEFNQGGSFIFATASMANIAQGAACLA	166
Db	1	YDFGGPVGGLLFGVYSPVITGLHQSFPPIELEFNQGGSFIFATASMANIAQGAACLA	60
QY	167	VFFIAKSEKUKGLAGASGSAVLGITTEPAIFGVNLRWRPFPIGTAAGGALIALFNI	226
Db	61	VFFIAKSEKUKGLAGASGSAVLGITTEPAIFGVNLRWRPFPIGTAAGGALIALFNI	120
QY	227	KAVAGAGFLGVVSDADPMVMFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAA	286
Db	121	KAVAGAGFLGVVSDADPMVMFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAA	180
QY	287	PVPAGTTKAEAEAPAEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQ	346
Db	181	PVPAGTTKAEAEAPAEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQ	240
QY	347	LVSFVSGKIIVVAPPFSGHAFVRTKAEDGNSVDILMHIGFDTVNLNGTHFNPLKKGDEVK	406
Db	241	LVSFVSGKIIVVAPPFSGHAFVRTKAEDGNSVDILMHIGFDTVNLNGTHFNPLKKGDEVK	300
QY	407	AGELLCEFDIDAIAKAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPA	466
Db	301	AGELLCEFDIDAIAKAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPA	360
QY	467	TP 468	
Db	361	TP 362	
RESULT 6			
ID	ADH87937 standard; protein; 496 AA.		
XX	ADH87937;		
DT	22-APR-2004 (first entry)		
DE	Enterococcus faecalis polypeptide #2417.		
XX	Enterococcus faecalis infection; transcription regulatory element;		
KW	antibacterial.		
OS	Enterococcus faecalis.		
XX	US6617156-B1.		
PN	09-SEP-2003.		
PD	13-AUG-1998; 98US-00134000.		
XX	15-AUG-1997; 97US-0055778P.		
XX	(DOUC/) DOUCETTE-STAMM L A.		
PA	(BUSH/) BUSH D.		
XX	Doucette-Stamm LA, Bush D;		
PI	WPI; 2003-895394/82.		
XX	N-PSDB; ADH84532.		
DR	New nucleic acid comprising a sequence encoding an Enterococcus faecalis		
XX	polypeptide, useful for preparing a composition for diagnosing or		
PT	treating E. faecalis infection.		
PS	Disclosure; SEQ ID NO 5822; 193pp; English.		
XX	The invention relates to Enterococcus faecalis polynucleotides and		
CC	polypeptides. The invention also relates to a recombinant expression		
CC	vector comprising a polynucleotide operably linked to a transcription		
CC	regulatory element, a cell comprising a recombinant vector, a method for		
CC	producing an E. faecalis polypeptide, an isolated nucleic acid comprising		
CC	a sequence not given in the specification, a recombinant vector		



ABU32951  
XX ABU32951 standard; protein; 617 AA.  
AC  
XX ABU32951;  
XX  
XX 19-JUN-2003 (first entry)  
DT  
XX Protein encoded by Prokaryotic essential gene #18478.  
DE  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
XX Listeria monocytogenes.  
XX  
XX WO20027183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
XX  
XX 06-SEP-2001; 2001US-00948993.  
XX  
XX 25-OCT-2001; 2001US-0342923P.  
XX  
XX 08-FEB-2002; 2002US-00072851.  
XX  
XX 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
XX  
XX N-PSDB; ACA36821.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 60875; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX proliferation; (7) identifying a compound that influences the activity of  
XX the gene product or that has an activity against a biological pathway;  
XX the gene product or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX the target prokaryotic essential genes. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 617 AA;  
SQ

Query Match 29.4%; Score 695; DB 6; Length 617;  
Best Local Similarity 34.8%; Pred. No. 6.7e-60;  
Matches 154; Conservative 82; Mismatches 162; Indels 44; Gaps 6;  
QY 2 AMVPPSLVNGYDVAATWAAGEMPWLSFLGLDVAAGYQGTVLPVLVVSWTILATIEKFLHK 61  
DB 185 ALVYPTMINLNFEGAHITFLQIP-----VLMYSYFSVPIIILAWFLSILIRFENS 236  
QY 62 RLKGTADFLITPVLTLLTGTFLTAIGPAMRWGDVLAHGLQGLYDFGPGVGGLLFLGLV 121  
DB 237 KIHEAAKTFTPMICMLIIVPLFLAFGLTFTFISQGLASGYTFIYNLSPIVAGAEWGAP 296  
QY 122 YSPVITGLHQSFPPIEL-ELFNQGGSFIFATASMANIAQAACLAFFFLAKSEKLGELA 180  
DB 297 WQVLVIRGHIHWGFVPIIMINLSRYGRDTMTAMWGPSNFAQAAGASLGVLTKTKPEVKAIA 356  
QY 181 GASVSAVLGITTEPAIFGVNLRWPFIFIGTAAIGGALIALFNKALGAAGFL--- 237  
DB 357 GSAALTGFGITPSIYGVTLKYKKPFVIASIAAGAIAGVGAAGSSGAANAIPIGLTLP 416  
QY 238 -----GVVSIDAPDMVMFLVCAVVTFFIAFGAALAYGLYLVRNNGSIDPDATAAFVAGT 292  
DB 417 IFIGKGFVGLGTAVAYIILSAIGTFYFGYKDEWADGI-----APT 457  
QY 293 TKAEAEAPAEFNSDSTTIQAPLTGEALIALSVSDAMFASGKLGSGVALVPTKGQIVSPVS 352  
DB 458 TKEAKETGVE---AEVIVSPIRGNIVPLNEVKDEAFSAGLLGRGVAIVPQEGKLISPVN 513  
QY 353 GKIIVAPSPGHAFVIRKAEKDSNVVDILMHIGFTDVLNLTGTHFNPLKKQGEVKAAGELLC 412  
DB 514 GTIETAPPTGHAIGIRS-----DKGVEILLHVGFDTVQLNGKIFKLLVAGQDRVLVQALL 569  
QY 413 EFDIDAIKAAGYEVTTPIVWSN 434  
DB 570 EFDLEAKADGYDTTPIVWVN 591  
RESULT 9  
ABU02378  
ID ABU02378 standard; protein; 655 AA.  
XX  
XX AC ABU02378;  
XX  
XX DT 23-OCT-2003 (revised)  
XX DT 11-FEB-2003 (first entry)  
XX  
XX S. pneumoniae type 4 strain protein from coding region #1956.  
XX  
XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
XX antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
XX gene therapy; vaccine.  
XX  
XX Streptococcus pneumoniae; type 4 strain.  
XX  
XX WO200277021-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 27-MAR-2002; 2002WO-IB002163.  
XX  
XX 27-MAR-2001; 2001GB-00007659.  
XX  
XX (CHIR-) CHIRON SPA.  
XX (GENO-) INST GENOMIC RES.  
XX  
XX Masignani V, Tettelin H, Fraser C;  
XX  
XX WPI; 2003-040579/03.  
XX  
XX N-PSDB; ABX07668.  
XX  
XX New proteins and nucleic acid molecules from *Streptococcus pneumoniae*,  
XX useful as medicaments for treating or preventing a disease or infection  
XX due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or



Db 371 HDEREQSLPAPISAYLGVTEPALGVNVKYYIPFVAGMTGSLAGMLSVTENVTAASI 430  
QY 232 GAAGFLGVSVSIDAPDMVMF-----LVCAVVTFFIAFGAIAIYGLYLVRRNGSIDPDATAAP 287  
Db 431 GIGGLPGLISIQPYMLPFAGTMLVAIVVPMLLTF-----FFRKAGLFTK----- 475  
QY 288 VPAGTTKAERAE-----APAEHSNDSTI-----TQAPLTGAEIALSSVSDAMFASGKLGG 337  
Db 476 -TEGDTNLOAEFVAQEAEEAEFVNEHPVELTSVEIISPLTGQVKELSQAIIDPFIFASGVWGQ 534  
QY 338 VATVPTKQGLVSVSGKIVVAFPSGHAFVARTKAEDGNSVDIIMHIGFDTVNLNGTHFNP 397  
Db 535 LVTEPSQGLTSPVNGTVTLFPTKHAIGI--VSDEG--VELLIHIGMDVGLDGKGFES 590  
QY 398 LKQGGDEVKAGELLCEFDIDAIKAAGYEVTTPVVSNN---YKKTGPNVTYGLGEIEAGAN 454  
Db 591 LVVQGDHVTVQQLIRFDMVDVKAAGLVTTETPVIITNQDAYTATIP-GTYPT-TIQAGAS 648  
QY 455 LL 456  
Db 649 LM 650  
RESULT 11  
ADK46609  
ID ADK46609 standard; protein; 627 AA.  
XX  
AC ADK46609;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Streptococcus pneumoniae protein, Seq ID No 3124.  
XX  
KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
XX  
OS Streptococcus pneumoniae.  
XX  
FN US6699703-B1.  
XX  
PD 02-MAR-2004.  
XX  
PF 26-MAY-2000; 2000US-00583110.  
XX  
PR 02-JUL-1997; 97US-0051553P.  
PR 12-MAY-1998; 98US-0085131P.  
PR 30-JUN-1998; 98US-00107433.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;  
XX  
XX WPI; 2004-212399/20.  
DR N-PSDB; ADK43948.  
XX  
XX New nucleic acid molecules and polypeptides useful for diagnosing,  
PT preventing and treating pathological conditions resulting from bacterial  
PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
PT screening.  
XX  
XX Disclosure; SEQ ID NO 3124; 301pp; English.  
XX  
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html.  
XX  
XX Sequence 627 AA;  
SQ

Query Match 28.3%; Score 669; DB 8; Length 627;  
Best Local Similarity 34.5%; Pred. No. 2,7e-57;  
Matches 150; Conservative 92; Mismatches 186; Indels 26; Gaps 10;  
QY 1 MAMVFSLVNGYDVAATMAAGEMPWNSLFLGLDVAQAGYQGTLPVLVWVSWILATIKFLH 60  
Db 182 MMLVSGSLPNMAVA--OQGEVTAMNFGF-IPVVLQGSVLPAFLIGVVGAKFEKAVR 237  
QY 61 KRLKGTADFLITPVLITLLTGTFTFAIGFAMRWGVDVLAHGLQGLYDFGCGPVGGLLFLG 120  
Db 238 KVVDPDVIDLLVTFVILLVMSILGLFVGPVHVVENYILIAATKAILSIPLGLGGFLIG 297  
QY 121 VYSPVITGLHQSPFPIELELNFQGSFIP-ATASHANIAQGAACLAFFFLAKSEKLKGL 179  
Db 298 VHQLIVVSGVHHIFNLLEVLQLAADHANPENAIITAAMTAQGAATVAVGVKTNPKUTL 357  
QY 180 AGASGVSAVLGITPEAIFGVNLRIRWPFPTIGTAAIGGALIALFNIAKAVAGAAFG- 238  
Db 358 APPAALSAGLITPEAIFGVNLRIRWPFPTIGTAAIGGALIALFNIAKAVAGAAFG- 417  
QY 239 VVSIDAPDMVFLVCAVVTFFIAFGAIAIYGLYLVRRNGSIDPDATAAPVPAGTTKAER 298  
Db 418 MLYVNGQLPQYLLMVAVSFALGFALTYMFGY-----EDEVDTAAAKQAEVAEKEE 470  
QY 299 -APAEFSNDSTIQAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKQLVSPVSGKIV 357  
Db 471 VAPALQNETLV--TPVGDVVALADVNDPVFSSGAMGQGIHAKPSQGVVTPADAEVSI 528  
QY 358 APPSGHAFVARTKAEDGNSVDIIMHIGFDTVNLNGTHFNPCLKQGDDEVKAGELLCEFDID 417  
Db 529 APTGTHAFGLKTR----NGAEVLHVGIDTVSMNGDGFGEAKVQGNKVGAGDVLGTFDSN 584  
QY 418 AIKAAGYEVTTPVVSNN---YKKTGPNVTYGLGEIEAGANLLNV 458  
Db 585 KIAAAGLDDTTMVTNTADYASVAPVAT---GSVAKGDAVIEV 625  
RESULT 12  
ABB49833  
ID ABB49833 standard; protein; 634 AA.  
XX  
AC ABB49833;  
XX  
DT 05-FEB-2002 (first entry)  
XX  
DE Listeria monocytogenes protein #2537.  
XX  
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
KW vitamin B12; bacterial infection; disease.  
XX  
OS Listeria monocytogenes.  
XX  
FN WO200177335-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 11-APR-2001; 2001WO-FR001118.  
XX  
PR 11-APR-2000; 2000FR-00004629.  
XX  
XX (INSP ) INST PASTEUR.  
XX  
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;  
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;  
XX  
XX WPI; 2002-010914/01.  
DR

XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment  
PT and prevention of *Listeria* and related bacterial infections, and related  
PT polypeptides.  
XX  
PS Claim 6; SEQ ID NO 2538; 192pp; French.  
XX  
CC The present invention relates to the genome sequence of *Listeria*  
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in *L.*  
CC monocytogenes and related organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of *L. monocytogenes* and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate *L. monocytogenes*-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccine compositions for the treatment or prevention of infections by *L.*  
CC monocytogenes and related organisms. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 634 AA;  
SQ  
  
Query Match 28.2%; Score 667; DB 5; Length 634;  
Best Local Similarity 34.6%; Pred. No. 4,3e-57;  
Matches 156; Conservative 81; Mismatches 172; Indels 42; Gaps 10;  
  
QY 2 AMVFPSEL--VNGYDVAATMAAG---EMPMSLFG-GLDVAQAQYQGTLPVLVVSMLATI 55  
DB 182 ALVYPTLAGITAGDPLYLTFAGTIFESPHVITFLGIPVILMSYASVPIIVATYFGSKV 241  
QY 56 EKFLHRLKGTADFILTPVLLTLLTGFTFTTIAIGPAMRWGVLDLAHGLQGLYDFGPGVG 115  
DB 242 EKGFKXIIIPDVIKTVPVFPCTLLIVVPTTIVIGPIATWAGQLGAGTIWYNLSPIIAG 301  
QY 116 LFLGLVYSPVITGLHQSPPEL-ELFNQGSFTFATASMANIAQGAACLVFLAKSE 174  
DB 302 LILGFWQVTVFLGLHWGLVPVAINNLTVLGHDPILAMTFGASFAQIGAVLAVFCKSRNK 361  
QY 175 KLKLAGASGVSAVLGITTEPAIFGVNLRWRPFFIGTIGTAAIGGALIALFNKAVALGAA 234  
DB 362 KIKSLISIPAFISGIPGVTETPAIYGVTLPLKXPFMINSIAGIGGILGFAGSQTYIMGL 421  
QY 235 GFLGVWSIDAP-----DMWFLVCVVTFTIAFGAATAYGL-----VLVRRNGSIDPDA 283  
DB 422 GIFGLPNEFKPGSGISGSEFWVWVAIVISFTLITVVGFKDPADVVVEQSNVTE--- 478  
QY 284 TAAPVPAGTTKAAEAPAEFSDNSTIIQAPLTGALALSSVSDAMFASGKLGSGVAIVPT 343  
DB 479 -----GETLIERET-----IPAPVVGHVITLADVKDEAFSGALGKGVAIIPT 521  
QY 344 KQQLVSPVSGKIVVAFPSGHAFVTRKAEQSDNVDILMHIGFTVNLNGTHFNPLKQGD 403  
DB 522 VGRVVAAGAVTFTIPTGHAIGITTK--DGA--EVLHIGNDIVQLEGKFTTAHVQGD 577  
QY 404 EVKAGELLCEFDIDAIKAAGVEVTPVIVSN 434  
DB 578 VIERGOLLTEFDIEGIIKAAGYDVTTPVVVTN 608  
  
RESULT 13  
ABP25654  
ID ABP25654 standard; protein; 676 AA.  
XX  
AC ABP25654;  
XX  
DT 02-JUL-2002 (first entry)  
XX

DE Streptococcus polypeptide SEQ ID NO 484.  
XX  
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX  
OS Streptococcus agalactiae.  
PN WO200234771-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 29-OCT-2001; 2001WO-GB004789.  
XX  
PR 27-OCT-2000; 2000GB-00026333.  
PR 24-NOV-2000; 2000GB-00028727.  
PR 07-MAR-2001; 2001GB-00005640.  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
PI Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;  
PI Tettelin H;  
XX  
DR WPI; 2002-352536/38.  
DR N-PSDB; ABN66285.  
XX  
XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.  
XX  
PS Claim 1; Page 3202-3203; 4525pp; English.  
XX  
CC The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins  
XX  
SQ Sequence 676 AA;  
  
Query Match 28.1%; Score 663.5; DB 5; Length 676;  
Best Local Similarity 35.3%; Pred. No. 1.1e-56;  
Matches 169; Conservative 82; Mismatches 195; Indels 33; Gaps 13;  
  
QY 1 MAMVFPSELVNGYDVAATMAAGEMPMSLFG-LDVAQAQYQGTLPVLVVSMLATI 58  
DB 212 ICLVSPQLLNAYSVASTSAADIKNWSWNGYFTVQKIGYQAQVIPALLAGLSLSYLEIF 271  
QY 59 LHKELKGTADFILTPVLLTLLTGFTFIATGP-----AMRWGVLDVAHGLQGLYDFGPGVG 114  
DB 272 WRKHPEVSMIFVFPFSLVPAITLAHTVLGIGTIGLTKWISAILVGLT-----GPVK 325  
QY 115 ---GLLFLGVSPVITGLHQSPPELLEFNQGSF---IFATASMANIAQGAACLVAF 168  
DB 326 WLFCAIFGALYAFVITGLHMTNAIDTLADTKHTTGLWPMIALSNIAQGSVAIVAY 385  
QY 169 FLAK-SEKLAGASGVSAVLGITTEPAIFGVNLRWRPFFIGTIGTAAIGGALIALFNK 227  
DB 386 FMHRHDEKQAQISLPAISAYLVGTEPALGVNVKVIYTPFVAGNIGSVAGLLATTENQ 445





**This Page Blank (uspto)**